CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN PARTICULAR FOR THE DIAGNOSIS, PREVENTION AND TREATMENT OF INFECTION

The subject of the invention is the genomic sequence and the nucleotide sequences encoding polypeptides of *Chlamydia trachomatis*, such as cellular envelope polypeptides, which are secreted or specific, or which are involved in metabolism, in the replication process or in virulence, as well as vectors including the said sequences and cells or animals transformed with these vectors. The invention also relates to methods of detecting these nucleic acids or polypeptides and kits for diagnosing *Chlamydia trachomatis* infection. The invention also relates to a method of selecting compounds capable of modulating bacterial infection and a method for the biosynthesis or biodegradation of molecules of interest using the said nucleotide sequences or the said polypeptides. The invention finally comprises, pharmaceutical, in particular vaccine, compositions for the prevention and/or treatment of bacterial, in particular *Chlamydia trachomatis*, infections.

The genus *Chlamydia* is composed of four species: *Chlamydia psittaci*, *Chlamydia pecorum*, *Chlamydia pneumoniae* and *Chlamydia trachomatis*.

- *Chlamydia psitacci* comprises numerous species, whose hosts are terrestrial vertebrate animals as well as birds and occasionally humans;
- *Chlamydia pecorum* is a pathogen of ruminants;
- *Chlamydia pneumoniae* is responsible for large pathies, for sinusites and for arteri in humans;
- *Chlamydia trachomatis* (Ct) is responsible for a number of human diseases:

- eye diseases: conjunctivitis in trachoma, paratrachoma, neonates and in adults;

- genital diseases: nongonococcal uretritis, epididymitis, cervicitis, salpingitis, perihepatitis and bartholinitis as well as pneumopathy in breast-feeding infants;
- systemic diseases: venereal lymphogranulomatosis (LGV).

These diseases affect a very large number of women and men [more than 600 million individuals are trachoma carriers and there are more than 90 million cases of genital *Chlamydia* infections] worldwide. Accordingly, basic and applied research which makes it possible to understand the physiopathology linked to this bacterium is very important for public health. (Raulston JE., 1995; Hackstadt T. et al., 1996).

Eye impairments due to *Chlamydia trachomatis* cause trachoma and inclusion conjunctivitis. Trachoma is a chronic conjunctivitis. It is the major cause of curable eye diseases leading to blindness. It is estimated that 20 million cases of loss of sight are due to it worldwide. Moreover, inclusion conjunctivitis is an eye inflammation which is caused by *Chlamydia trachomatis* and is transmitted by the venereal route. Inclusion conjunctivitis affects adults and neonates exposed to genital secretions.

Two types of eye disease caused by agents of the species *Chlamydia trachomatis* can be distinguished. The conventional trachomatous disease is found in endemic regions; transmission occurs from eye to eye and through the hands, or it can be passed on by flies In nonendemic regions, transmission occurs through genital apparatus; it usually only ed conjunctivitis, most often without as lar keratitis; it is rare for a pannus or for sca tival to those in trachoma to develop. This ate it impairment is called paratrachoma to di ich is from the conventional endemic tra ness and transmitted by the ocular route. d over the the number of cases of trachoma h mprovement in last forty years. This is relat

hygiene and living conditions. However, trachoma remains the principal cause of avoidable blindness in Africa, in the Middle East and in some regions of Asia. The transmission of the endemic disease occurs in particular through close personal contact, in regions where a secondary exposure exists in a repeated form. Often, the infection is also latent. In some industrialized countries, such as the United States, a mild form of trachoma still exists in some ethnic groups. Sometimes, a tardive trachoma may be found following an immunosuppressive treatment. The eye impairments caused by *Chlamydia trachomatis*, such as inclusion conjunctivitis and paratrachoma, are also a complication due to a common venereal infection. These infections are not very frequent; they occur most often in young adults. The eye impairments in neonates are produced during the passage through the maternal genital routes during childbirth. Theoretically, endemic trachoma and inclusion conjunctivitis in adults appear in the form of conjunctivitis, the latter being characterized by the presence of lymphoid follicles. In regions where the endemic disease is serious, the disease often starts before the age of 2 years and reinfection is frequent. Superficial neovascularization is added, in this case, to leukocytic infiltration. The conjunctival scars will then cause trichiasis and entropion. The eroded cornea will become a carrier of a corneal ulcer of bacterial origin. The scar on the cornea causes blindness. Impairment of the lachrymal glands gives a picture of dryness of the cornea. Xerosis becomes complicated with secondary bacterial ulcer. In regions where trachoma is endemic, the infectious process disappears towards the age of fifteen. The scars then progress to blindness, which affects almost exclusively adults. In regions where exposure is lower, the infectious process is, in this case, less rapid and adults are carriers of a chronic disease.

Positive diagnosis of trachoma can be most

RWS Group plc, of Europa House, Marsham Way, Gerrards Cross, Buckinghamshire, England, hereby declares that, to the best of its knowledge and belief, the following document, prepared by one of its translators competent in the art and conversant with the English and French languages, is a true and correct translation of the accompanying documents in the French language.

Signed this 19th day of November 1998

S. DRANE

Director

For and on behalf of RWS Group plc

often established by clinical observation: lymphoid follicles are visible on the upper tarsal conjuctiva; conjunctival scar is typical. Vascular pannus exists. In endemic regions, clinical diagnosis is often sufficient. However, isolated cases of inclusion conjunctivitis must be the subject of a differential diagnosis, in particular to distinguish viral conjunctivitis.

Public health measures against the endemic form of the disease provide for mass treatments with tetracycline or erythromycin collyria of all children. The treatment may also provide for surgical correction of the lesions. The other conjunctival impairments respond well to general treatments with tetracyclines or erythromycin. The prevention of trachomatous disease by health measures and by improving living standards is sufficient. Furthermore, to avoid the spread of trachoma, antibiotic collyria may be used.

The role of *Chlamydia trachomatis* in a number of genital impairments has been demonstrated over the last three decades. *Chlamydia trachomatis* is responsible in this case for a pathology which may be superposed on the impairments observed with *Neisseria gonorrhoeae*. The pathologies for which *Chlamydia trachomatis* may be responsible at the genital level are acquired by the venereal route and are a major source of sexually transmitted diseases.

The epidemiology of *Chlamydia trachomatis* genital infections shows each year more than 4 million new cases in the United States, and more than 3 million new cases in Europe. Like the other venereal infections, *Chlamydia trachomatis* affects young subjects. There is a direct relationship between the number of sexual partners and the frequency of the disease. For example, the frequency of *Chlamydia trachomatis* appears to be five to ten times higher than that of *Neisseria gonorrhoeae* in pregnant women. The *Chlamydia trachomatis* infection is probably more discreet than its *Neisseria gonorrhoeae* homologue. This

relative clinical silence, estimated in women at 50% or even 70% of infections, explains why the total morbidity of *Chlamydia trachomatis* conditions is high. Diagnosis must therefore be requested in patients who are sometimes asymptomatic carriers of infection.

Chlamydia trachomatis is responsible for nearly 30% of nongonococcal urethritis, or NGU. *Chlamydia trachomatis* urethritis may be discreet, the disease then progresses to a certain form of chronicity. The diagnosis will, like for the other clinical forms of the disease, be called into play later.

Chlamydia trachomatis is a cause of epididymitis in humans during a period of sexual activity. The bacterium may be found in the urethra, urine, sperm or even a sample collected by aspiration from the epididymis. It is in particular found in humans under 35 years of age. A discharge from the urethra which is associated with the disease suggests the diagnosis of a *Chlamydia* condition or sometimes a gonococcal condition.

Untreated Reiter's syndrome, if accompanied by urethritis, evokes a *Chlamydia trachomatis* condition.

Chlamydia trachomatis affects 30% to 40% of women who are clinically carriers of a gonorrhoea (or have had contact), 10% to 20% of women having a venereal origin, 5% of women consulting having no particular origin.

The cervix is often normal during a *Chlamydia trachomatis* infection. However, a hypertrophic cervical erythema will cause such an infection to be suspected. *Chlamydia trachomatis* is responsible for an endocervicitis whereas viral impairments result in exocervicitis. A nongonococcal endocervicitis requires treating the patient and partners with tetracyclines.

Chlamydia trachomatis is responsible for a large number of acute salpingites. The picture is often complicated by an acute peritonitis or even a perihepatitis.

In case of pregnancy, the risk is first that of

infection of the neonate at birth. However, the risk of postpartum complications exists (endometritis or salpingitis).

The reference method for the diagnosis of *Chlamydia trachomatis* is the isolation of the bacterium on cell culture. For all infections, the sample collection should make it possible to obtain a suitable sample with the aid of a swab. This sample should be transported to a laboratory under excellent conditions; in particular, the cold chain must absolutely be maintained. The placing in cell culture on mouse fibroblasts will be carried out by people having specific skills. The distinction of *Chlamydia trachomatis* with labelled antibodies and the observation of cell cultures under a microscope will take place two days after placing in culture. Provided these imperatives are observed, cell culture is a reliable technique. However, the constraints linked to this technique are many: not only must the laboratory be equipped for the cell culture, but, furthermore, highly competent staff must take care of this type of diagnosis.

Techniques for identifying genetic material can obviously be used for the detection of *Chlamydia trachomatis*. Among these techniques, enzymatic gene amplification or PCR is favoured by those skilled in the art. The technique indeed makes it possible to identify *Chlamydia trachomatis* with a very high sensitivity and complete specificity. Initially used in specialist laboratories, PCR is now performed in numerous medical laboratories. This diagnostic approach is important because it allows detection of the bacteria even in samples which have been transported under poor conditions.

The treatment of *Chlamydia* urethritis with antibiotics such as tetracycline or quinolones is very effective. The duration of treatment varies between 7 and 14 days. The treatment of pregnant women poses the problem of contraindications to tetracycline.

Neonatal infections caused by *Chlamydia trachomatis* are explained by the frequency of these bacteria in the cervix. In some studies, 5% to 13% of impairments are observed in the cervix in asymptomatic pregnant women. The neonates risk, in this case, developing an inclusion conjunctivitis. Not only can *Chlamydia trachomatis* be isolated from the children's eyes, but also persistently from the rhinopharynx and also from the rectum. Pneumopathies and otitis media are also found, a result of contamination at childbirth.

Differential diagnosis of inclusion conjunctivitis in neonates is required with gonococcal ophthalmia; while the duration of incubation is from one to three days in the case of a gonococcal ophthalmia, neonatal inclusion conjunctivitis has an acute beginning with discharge and formation of membranes or even of conjunctival scars.

Treatment consists of oral erythromycin at the dose of 40 to 50 mg per kg of weight, for two to three weeks. In a nonendemic trachoma region, this disease never progresses to chronicity.

Finally, mention should be made of infantile pneumopathy. The syndrome is well defined; it is found in children affected by *Chlamydia trachomatis*. Less than ten children are affected by *Chlamydia trachomatis* pneumopathies per thousand births. The syndrome is, in this case, always found at an early age (less than four months).

Venereal lymphogranulomatosis is an infection which is transmitted through sexual contact and is due to *Chlamydia trachomatis* strains L1, L2 and L3. In humans, a passing primary genital lesion is followed by an often suppurative and multiple regional lymphadenopathy. This disease is a general disease which is accompanied by fever and a rise in the number of white blood cells. If it progresses to chronicity, the disease then becomes complicated with genital elephantiasis, stricture or even fistula of the genital

apparatus, of the penis, of the urethra and of the rectum.

The three *Chlamydia trachomatis* strains L1, L2 and L3 are responsible for venereal lymphogranulomatosis. These *Chlamydia* strains are more virulent than the strains responsible for trachoma and STD. It is very important to note that venereal lymphogranulomatosis is a systemic disease which affects primarily the lymphatic tissue. Generally transmitted by the sexual route, *Chlamydia trachomatis* L may also cause contamination through direct contact or even during poor laboratory handling. In spite of these variable modes of transmission, the age for the highest incidence of these diseases corresponds to that for greater sexual activity. Venereal lymphogranulomatosis is still endemic in South America, in Africa and sometimes in Asia. For a long time, the prevalence of venereal lymphogranulomatosis was difficult to establish because of the difficulty of performing diagnosis with certitude. It should also be noted that men are affected more often than women. In low endemic regions, it is difficult to recognize the reservoir of microbes. This situation is explained by the fact that the isolation of the strains causing venereal lymphogranulomatoses from asymptomatic subjects is rarely successful.

Clinical impairment by venereal lymphogranulomatosis manifests itself by the appearance of a small ulcer 3 to 21 days after the exposure of small nonpainful vesicles. In both men and women, the lesion is most often silent. Since this impairment disappears within a few days and causes no functional discomfort and leaves no visible scar, the disease is often recognized late. The venereal lymphogranulomatosis strains may be found in the urethra or the endocervix in patients with inguinal adenopathies; these regions are then considered as the initial site of infection. The characteristic feature of the venereal lymphogranulomatosis strains is that from the initial site of

infection, *Chlamydia* exhibits a diffusion drained by the lymphatic ducts. The disease is then complicated by a ganglionic impairment of the region draining the site of inoculation. By way of example, anorectal infection causes deep adenopathies. These adneopathies are marked by the appearance of a periadenitis which forms a fluctuating and suppurative ganglionic mass. Fistulae will appear during the decline of the disease. As general signs are present at this stage of the disease, it is often confused with a malignant lymphoma. The other general complications are rarely observed. Clinical examinations have been able to lead biologists to isolate *Chlamydia* from the cerebrospinal fluid or from the blood. It should also be noted that in a number of cases (5%), venereal lymphogranulomatosis is complicated by a chronic oedema: this is genital elephantiasis.

The diagnosis of venereal lymphogranulomatosis requires the isolation of the *Chlamydia* strains involved in the disease. However, isolation on cell cultures is rarely used, but immunological reactions may be used.

The treatment of venereal lymphogranulomatosis in its initial phase is identical to the treatment of other *Chlamydia* infections. In the chronic phases, antibiotics have little effect on the progress of the disease, but they are however useful in case of super-infection. Although the recommended therapeutic arsenal is identical, it is advisable to prolong the treatment for a period of at least four weeks. In addition to this treatment, reconstructive surgery may be useful in cases of urethral, penile or rectal strictures, as well as for the treatment of fistulae.

In conclusion, a short and effective treatment, without recurrences, and a well-tolerated treatment of *Chlamydia trachomatis* infections therefore remains desirable.

An even greater need up until now relates to a diagnosis which is specific to each of the strains,

which is sensitive, which can be carried out conveniently and rapidly, and which allows early detection of the infection.

No vaccine is currently available against *Chlamydia trachomatis*. The role of the immune defence in the physiology and pathology of the disease should probably be understood in order to develop satisfactory vaccines.

More detailed information relating to the biology of these strains, their interactions with their hosts, the associated phenomena of infectivity and those of escaping the immune defences of the host in particular, and finally their involvement in the development of the these associated pathologies, will allow a better understanding of these mechanisms. In the light of the preceding text which shows in particular the limitations of the means of controlling *Chlamydia trachomatis* infection, it is therefore at present essential, on the one hand, to develop molecular tools, in particular from a better genetic knowledge of *Chlamydia trachomatis*, but also to develop new preventive and therapeutic treatments, new diagnostic methods and new vaccine strategies which are specific, effective and tolerated. This is precisely the object of the present invention.

The subject of the present invention is the nucleotide sequence having the sequence SEQ ID No. 1 of the *Chlamydia trachomatis LGV2* genome.

The subject of the present invention is also nucleotide sequences characterized in that they are chosen from:

a) a nucleotide sequence exhibiting at least 99.9% identity with the sequence SEQ ID No. 1;

b) a nucleotide sequence homologous to the sequence SEQ ID No. 1;

c) a nucleotide sequence complementary to the sequence SEQ ID No. 1 or complementary to a nucleotide sequence as defined in a) or b), and a nucleotide sequence of their corresponding RNA;

d) a nucleotide sequence of a representative fragment of the sequence SEQ ID No. 1, or of a representative fragment of the nucleotide sequence as defined in a), b) or c);

e) a nucleotide sequence comprising a sequence as defined in a), b), c) or d);

f) a nucleotide sequence capable of being obtained from a nucleotide sequence as defined in a), b), c), d) or e); and

g) a modified nucleotide sequence of a nucleotide sequence as defined in a), b), c), d), e) or f).

Sequence of the genome, or genomic sequence of *Chlamydia trachomatis* is understood to mean the sequence of the chromosome of *Chlamydia trachomatis*, in contrast with the plasmid sequence of *Chlamydia trachomatis*.

Nucleotide sequence, polynucleotide or nucleic acid are understood to mean, according to the present invention, either a double-stranded DNA, a single-stranded DNA or products of transcription of the said DNAs.

It should be understood that the present invention does not relate to the genomic nucleotide sequences of *Chlamydia trachomatis* taken in their natural environment, that is to say in the natural state. They are sequences which may have been isolated, purified or partially purified, by separation methods such as, for example, ion-exchange chromatography, molecular size exclusion chromatography or affinity chromatography, or alternatively fractionation techniques based on solubility in various solvents, or by genetic engineering methods such as amplification, cloning or subcloning, it being possible for the sequences of the invention to be carried by vectors.

The nucleotide sequence SEQ ID No. 1 was obtained by sequencing the *Chlamydia trachomatis LGV2* genome by the method of directed sequencing after fluorescent automated sequencing of the inserts of clones and assembling of these sequences of nucleotide

fragments (inserts) by means of softwares (cf. Examples). In spite of the high precision of the sequence SEQ ID No. 1, it is possible that it does not perfectly, 100% represent the nucleotide sequence of the *Ct LGV2* genome and that a few rare sequencing errors or uncertainties still remain in the sequence SEQ ID No. 1. In the present invention, the presence of an uncertainty for an amino acid is designated by "Xaa" and that for a nucleotide is designated by "N" in the sequence listing below. These few rare errors or uncertainties could be easily detected and corrected by persons skilled in the art using the entire chromosome and/or its representative fragments according to the invention and standard amplification, cloning and sequencing methods, it being possible for the sequences obtained to be easily compared, in particular by means of a computer software and using computer-readable media for recording the sequences according to the invention as described, for example, below. After correcting these possible rare errors or uncertainties, the corrected nucleotide sequence obtained would still exhibit at least 99.9% identity with the sequence SEQ ID No. 1.

Homologous nucleotide sequence for the purposes of the present invention is understood to mean a nucleotide sequence having a percentage identity with the bases of the nucleotide sequence SEQ ID No. 1 of at least 80%, preferably 90% and 95%, this percentage being purely statistical and it being possible for the differences between the two nucleotide sequences to be distributed randomly and over their entire length. The said homologous sequences exhibiting a percentage identity with the bases of the nucleotide sequence SEQ ID No. 1 of at least 80%, preferably 90% and 95%, may comprise, for example, the sequences corresponding to the genomic sequence or to the sequences of its representative fragments of a bacterium belonging to the Chlamydia family, including the species *Chlamydia pneumoniae*, *Chlamydia psittaci* and *Chlamydia pecorum*

mentioned above, as well as the sequences corresponding to the genomic sequence or to the sequences of its representative fragments of a bacterium belonging to the variants of the species *Chlamydia trachomatis*. In the present invention, the terms family and genus are mutually interchangeable, the terms variant, serotype, strain and subspecies are also mutually interchangeable. These homologous sequences may thus correspond to variations linked to mutations within the same species or between species and may correspond in particular to truncations, substitutions, deletions and/or additions of at least one nucleotide. The said homologous sequences may also correspond to variations linked to the degeneracy of the genetic code or to a bias in the genetic code which is specific to the family, to the species or to the variant and which are likely to be present in *Chlamydia*.

Nucleotide sequence complementary to a sequence of the invention is understood to mean any DNA whose nucleotides are complementary to those of the sequence of the invention, and whose orientation is reversed (antiparallel sequence).

Representative fragments of the sequences according to the invention will be understood to mean any nucleotide fragment having at least 8 successive nucleotides, preferably at least 12 successive nucleotides, and still more preferably at least 20 successive nucleotides of the sequence from which it is derived.

Among these representative fragments, those capable of hybridizing under stringent conditions with a nucleotide sequence according to the invention are preferred. Hybridization under stringent conditions means that the temperature and ionic strength conditions are chosen such that they allow hybridization to be maintained between two complementary DNA fragments.

By way of illustration, high stringency conditions for the hybridization step for the purposes

of defining the nucleotide fragments described above, are advantageously the following.

The hybridization is carried out at a preferred temperature of 65°C in the presence of SSC buffer, 1 × SSC corresponding to 0.15 M NaCl and 0.05 M Na citrate. The washing steps may be, for example, the following:

- 2 × SSC, 0.1% SDS at room temperature followed by three washes with 1 × SSC, 0.1% SDS; 0.5 × SSC, 0.1% SDS; 0.1 × SSC, 0.1% SDS at 68°C for 15 minutes.

Intermediate stringency conditions, using, for example, a temperature of 60°C in the presence of a 5 × SSC buffer, or of low stringency, for example a temperature of 50°C in the presence of a 5 × SSC buffer, respectively require a lower overall complementarity for the hybridization between the two sequences.

The stringent hybridization conditions described above for a polynucleotide of about 300 bases in size will be adapted by persons skilled in the art for larger- or smaller-sized oligonucleotides, according to the teaching of Sambrook et al., 1989.

Among the representative fragments according to the invention, those which can be used as primer or probe in methods which make it possible to obtain homologous sequences or their fragments according to the invention, or to reconstitute a genomic fragment found to be incomplete in the sequence SEQ ID No. 1 or carrying an error or an uncertainty, are also preferred, these methods, such as the polymerase chain reaction (PCR), cloning and sequencing of nucleic acid being well known to persons skilled in the art. These homologous nucleotide sequences corresponding to mutations or to inter- or intra-species variations, as well as the complete genomic sequence or one of its fragments capable of being reconstituted, of course form part of the invention.

Among the said representative fragments, those which can be used as primer or probe in methods allow-

ing diagnosis of the presence of *Chlamydia trachomatis* or one of its associated microorganisms as defined below are also preferred.

The representative fragments capable of modulating, regulating, inhibiting or inducing the expression of a gene of *Chlamydia trachomatis* or one of its associated microorganisms, and/or capable of modulating the replication cycle of *Chlamydia trachomatis* or one of its associated microorganisms in the host cell and/or organism, are also preferred. Replication cycle is intended to designate invasion, multiplication, intracellular localization, in particular retention in the vacuole and inhibition of the process of fusion to the lysosome, and propagation of *Chlamydia trachomatis* or one of its associated microorganisms from host cells to host cells.

Among the said representative fragments, those corresponding to nucleotide sequences corresponding to open reading frames, called ORF sequences (ORF for open reading frame), and encoding polypeptides, such as for example, but without being limited thereto, the ORF sequences which will be later described, are finally preferred.

The representative fragments according to the invention may be obtained, for example, by specific amplification, such as PCR, or after digestion, with appropriate restriction enzymes, of nucleotide sequences according to the invention; these methods are in particular described in the manual by Sambrook et al., 1989. The said representative fragments may also be obtained by chemical synthesis when they are not too large in size and according to methods well known to persons skilled in the art.

The representative fragments according to the invention may be used, for example, as primer, to reconstitute some of the said representative fragments, in particular those in which a portion of the sequence is likely to be missing or imperfect, by methods well known to persons skilled in the art such as amplifi-

cation, cloning or sequencing techniques.

Modified nucleotide sequence will be understood to mean any nucleotide sequence obtained by mutagenesis according to techniques well known to persons skilled in the art, and exhibiting modifications in relation to the normal sequences, for example mutations in the regulatory and/or promoter sequences for the expression of a polypeptide, in particular leading to a modification of the level of expression of the said polypeptide or to a modulation of the replicative cycle.

Modified nucleotide sequence will also be understood to mean any nucleotide sequence encoding a modified polypeptide as defined below.

The subject of the present invention is *Chlamydia trachomatis* nucleotide sequences characterized in that they are chosen from the ORF2 to ORF1076 sequences.

The ORF2 to ORF1076 nucleotide sequences are defined in Table 1 represented below by their position on the sequence SEQ ID No. 1. For example, the ORF10 sequence is defined by the nucleotide sequence between the nucleotides at position 9828 and 10430 on the sequence SEQ ID No. 1, ends included.

The invention also relates to the nucleotide sequences characterized in that they comprise a nucleotide sequence chosen from:

a) an ORF2 to ORF1076 nucleotide sequence according to the invention;

b) a homologous nucleotide sequence exhibiting at least 80% identity with an ORF2 to ORF1076 nucleotide sequence according to the invention or as defined in a);

c) a complementary or RNA nucleotide sequence corresponding to an ORF2 to ORF1076 sequence according to the invention or as defined in a) or b);

d) a nucleotide sequence of a representative fragment of an ORF2 to ORF1076 sequence according to the invention or of a sequence as defined in a), b) or c);

e) a nucleotide sequence capable of being obtained

from an ORF2 to ORF1076 sequence according to the invention or as defined in a), b), c) or d); and

f) a modified nucleotide sequence of an ORF2 to ORF1076 sequence according to the invention or as defined in a), b), c), d) or e).

As regards the homology with the ORF2 to ORF1076 nucleotide sequences, the homologous sequences exhibiting a percentage identity with the bases of one of the ORF2 to ORF1076 nucleotide sequences of at least 80%, preferably 90% and 95%, are preferred. The said homologous sequences correspond to the homologous sequences as defined above and may comprise, for example, the sequences corresponding to the ORF sequences of a bacterium belonging to the Chlamydia family, including the species *Chlamydia pneumoniae*, *Chlamydia psittaci* and *Chlamydia pecorum* mentioned above, as well as the sequences corresponding to the ORF sequences of a bacterium belonging to the variants of the species *Chlamydia trachomatis*. These homologous sequences may likewise correspond to variations linked to mutations within the same species or between species and may correspond in particular to truncations, substitutions, deletions and/or additions of at least one nucleotide. The said homologous sequences may also correspond to variations linked to the degeneracy of the genetic code or to a bias in the genetic code which is specific to the family, to the species or to the variant and which are likely to be present in *Chlamydia*.

The invention comprises the polypeptides encoded by a nucleotide sequence according to the invention, preferably by a representative fragment of the sequence SEQ ID No. 1 and corresponding to an ORF sequence, in particular the *Chlamydia trachomatis* polypeptides, characterized in that they are chosen from the sequences SEQ ID No. 2 to SEQ ID No. 1076.

The invention also comprises the polypeptides characterized in that they comprise a polypeptide chosen from:

a) a polypeptide according to the invention;

b) a polypeptide homologous to a polypeptide according to the invention, or as defined in a);

c) a fragment of at least 5 amino acids of a polypeptide according to the invention, or as defined in a) or b);

d) a biologically active fragment of a polypeptide according to the invention, or as defined in a), b) or c); and

e) a modified polypeptide of a polypeptide according to the invention, as defined in a), b), c) or d).

In the present description, the terms polypeptide, peptide and protein are interchangeable.

It should be understood that the invention does not relate to the polypeptides in natural form, that is to say that they are not taken in their natural environment but that they may have been isolated or obtained by purification from natural sources, or alternatively obtained by genetic recombination, or else by chemical synthesis and that they may, in this case, comprise nonnatural amino acids, as will be described below.

Homologous polypeptide will be understood to designate the polypeptides exhibiting, in relation to the natural polypeptide, certain modifications such as in particular a deletion, addition or substitution of at least one amino acid, a truncation, an extension, a chimeric fusion, and/or a mutation, or polypeptides exhibiting post-translational modifications. Among the homologous polypeptides, those whose amino acid sequence exhibits at least 80%, preferably 90%, homology with the amino acid sequences of the polypeptides according to the invention are preferred. In the case of a substitution, one or more consecutive or nonconsecutive amino acids are replaced by "equivalent" amino acids. The expression "equivalent" amino acid is intended here to designate any amino acid capable of being substituted for one of the amino acids in the basic structure without, however, essentially

modifying the biological activities of the corresponding peptides and as will be defined later.

These equivalent amino acids may be determined either based on their structural homology with the amino acids for which they are substituted, or on results of comparative tests of biological activity between the various polypeptides which may be carried out.

By way of example, there may be mentioned the possibilities of substitutions which may be carried out without resulting in a substantial modification of the biological activity of the corresponding modified polypeptides; the replacements, for example, of leucine with valine or isoleucine, of aspartic acid with glutamic acid, of glutamine with asparagine, of arginine with lysine, and the like, the reverse substitutions naturally being feasible under the same conditions.

The homologous polypeptides also correspond to the polypeptides encoded by the homologous nucleotide sequences as defined above and thus comprise in the present definition the mutated polypeptides or polypeptides corresponding to inter- or intra-species variations which may exist in *Chlamydia*, and which correspond in particular to truncations, substitutions, deletions and/or additions of at least one amino acid residue.

Biologically active fragment of a polypeptide according to the invention will be understood to designate in particular a polypeptide fragment, as defined below, exhibiting at least one of the characteristics of the polypeptides according to the invention, in particular in that it is:

- capable of eliciting an immune reaction directed against *Chlamydia trachomatis*; and/or
- capable of being recognized by an antibody specific for a polypeptide according to the invention; and/or
- capable of binding to a polypeptide or to a

nucleotide sequence of *Chlamydia trachomatis*; and/or

- capable of modulating, regulating, inducing or inhibiting the expression of a gene of *Chlamydia trachomatis* or one of its associated microorganisms, and/or capable of modulating the replication cycle of *Chlamydia trachomatis* or one of its associated microorganisms in the host cell and/or organism; and/or
- capable of generally exerting an even partial physiological activity, such as for example a structural activity (cellular envelope, ribosome), an enzymatic (metabolic) activity, a transport activity, an activity in the secretion or in the virulence.

Polypeptide fragment according to the invention is understood to designate a polypeptide comprising a minimum of 5 amino acids, preferably 10 amino acids and 15 amino acids.

The polypeptide fragments according to the invention may correspond to isolated or purified fragments which are naturally present in *Chlamydia trachomatis* or which are secreted by *Chlamydia trachomatis,* or may correspond to fragments capable of being obtained by cleaving the said polypeptide with a proteolytic enzyme, such as trypsin or chymotrypsin or collagenase, or with a chemical reagent, such as cyanogen bromide (CNBr) or alternatively by placing the said polypeptide in a highly acidic environment, for example at pH 2.5. Such polypeptide fragments may be equally well prepared by chemical synthesis, using hosts transformed with an expression vector according to the invention containing a nucleic acid allowing the expression of the said fragments, placed under the control of appropriate elements for regulation and/or expression.

"Modified polypeptide" of a polypeptide according to the invention is understood to designate a polypeptide obtained by genetic recombination or by chemical synthesis as will be described below, exhibiting at least one modification in relation to the normal sequence. These modifications may in particular

affect amino acids responsible for a specificity or for the efficiency of the activity, or responsible for the structural conformation, for the charge or for the hydrophobicity, and for the capacity for multimerization and for membrane insertion of the polypeptide according to the invention. It is thus possible to create polypeptides with an equivalent, an increased or a reduced activity, and with an equivalent, a narrower or a broader specificity. Among the modified polypeptides, there may be mentioned the polypeptides in which up to 5 amino acids may be modified, truncated at the N- or C-terminal end, or alternatively deleted, or else added.

As is indicated, the modifications of the polypeptide may have in particular the objective:

- of making it capable of modulating, regulating, inhibiting or inducing the expression of a gene of *Chlamydia*, in particular of *Chlamydia trachomatis* and its variants, or one of its associated microorganisms, and/or capable of modulating the replication cycle of *Chlamydia*, in particular of *Chlamydia trachomatis* and its variants, or one of its associated microorganisms, in the host cell and/or organism,
- of allowing its use in methods of biosynthesis or of biodegradation, or its incorporation into vaccine compositions,
- of modifying its bioavailability as a compound for therapeutic use.

The said modified polypeptides may also be used on any cell or microorganism for which the said modified polypeptides will be capable of modulating, regulating, inhibiting or inducing gene expression, or of modulating the growth or the replication cycle of the said cell or of the said microorganism. The methods allowing demonstration of the said modulations on eukaryotic or prokaryotic cells are well known to persons skilled in the art. The said cells or microorganisms will be chosen, in particular, from tumour cells or infectious microorganisms and the said

modified polypeptides may be used for the prevention or treatment of pathologies linked to the presence of the said cells or of the said microorganisms. It is also clearly understood that the nucleotide sequences encoding the said modified polypeptides may be used for the said modulations, for example by the intermediacy of vectors according to the invention and which are described below, so as to prevent or to treat the said pathologies.

The above modified polypeptides may be obtained using combinatory chemistry, in which it is possible to systematically vary portions of the polypeptide before testing them on models, cell cultures or microorganisms for example, so as to select the compounds which are the most active or which exhibit the desired properties.

Chemical synthesis also has the advantage of being able to use:

- nonnatural amino acids, or
- nonpeptide bonds.

Accordingly, in order to extend the life of the polypeptides according to the invention, it may be advantageous to use nonnatural amino acids, for example in the D form, or alternatively amino acid analogues, in particular sulphur-containing forms for example.

Finally, the structure of the polypeptides according to the invention, its homologous or modified forms, as well as the corresponding fragments may be integrated into chemical structures of the polypeptide type and the like. Accordingly, it may be advantageous to provide at the N- and C-terminal ends compounds which are not recognized by proteases.

Also forming part of the invention are the nucleotide sequences encoding a polypeptide according to the invention.

More particularly, the subject of the invention is nucleotide sequences, characterized in that they encode a polypeptide of the cellular envelope, preferably of the outer cellular envelope of *Chlamydia*

trachomatis or one of its fragments, such as for example the predominant proteins of the outer membrane, the adhesion proteins or the proteins entering into the composition of the Chlamydia wall. Among these sequences, the sequences comprising a nucleotide sequence chosen from the following sequences are most preferred:

ORF3; ORF19; ORF51; ORF189; ORF212; ORF213; ORF324; ORF477; ORF478; ORF479; ORF481; ORF482; ORF483; ORF484; ORF486; ORF488; ORF489; ORF490; ORF572; ORF573; ORF742; ORF817; ORF818; ORF820; ORF1035; ORF1036; ORF1037; ORF1038; ORF1070; ORF1071; ORF1073 and one of their representative fragments.

The structure of the cytoplasmic membranes and of the wall of bacteria is dependent on the associated proteins. The structure of the cytoplasmic membrane makes it impermeable to water, to water-soluble substances and to small-sized molecules (ions, small inorganic molecules, peptides or proteins). To enter into or to interfere with a cell or a bacterium, a ligand must establish a special relationship with a protein anchored in the cytoplasmic membrane (the receptor). These proteins which are anchored on the membrane play an important role in metabolism since they control the exchanges in the bacterium. These exchanges apply to molecules of interest for the bacterium (small molecules such as sugars and small peptides) as well as undesirable molecules for the bacterium such as antibiotics or heavy metals.

The double lipid layer structure of the membrane requires the proteins which are inserted therein to have hydrophobic domains of about twenty amino acids forming an alpha helix. Predominantly hydrophobic and potentially transmembrane regions may be predicted from the primary sequence of the proteins, itself deduced from the nucleotide sequence. The presence of one or more putative transmembrane domains raises the possibility for a protein to be associated with the cytoplasmic membrane and to be able to play an

important metabolic role therein or alternatively for the protein thus exposed to be able to exhibit potentially protective epitopes.

If the proteins inserted into the membrane exhibit several transmembrane domains capable of interacting with one another via electrostatic bonds, it then becomes possible for these proteins to form pores which go across the membrane which becomes permeable for a number of substances. It should be noted that proteins which do not have transmembrane domains may also be anchored by the intermediacy of fatty acids in the cytoplasmic membrane, it being possible for the breaking of the bond between the protein and its anchor in some cases to be responsible for the release of the peptide outside the bacterium.

Preferably, the invention relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 1 and 3 transmembrane domains and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF2; ORF3; ORF5; ORF8; ORF9; ORF10; ORF11; ORF12; ORF17; ORF21; ORF26; ORF27; ORF28; ORF29; ORF30; ORF31; ORF33; ORF35; ORF37; ORF39; ORF40; ORF41; ORF42; ORF43; ORF44; ORF45; ORF46; ORF47; ORF48; ORF49; ORF52; ORF53; ORF55; ORF56; ORF58; ORF65; ORF66; ORF68; ORF70; ORF74; ORF75; ORF76; ORF78; ORF79; ORF81; ORF82; ORF83; ORF86; ORF91; ORF92; ORF94; ORF97; ORF100; ORF102; ORF103; ORF105; ORF106; ORF107; ORF109; ORF110; ORF111; ORF112; ORF113; ORF114; ORF115; ORF116; ORF117; ORF120; ORF122; ORF123; ORF130; ORF134; ORF135; ORF137; ORF140; ORF141; ORF143; ORF144; ORF145; ORF147; ORF148; ORF149; ORF150; ORF151; ORF155; ORF156; ORF162; ORF163; ORF164; ORF165; ORF166; ORF167; ORF168; ORF169; ORF170; ORF171; ORF173; ORF175; ORF176; ORF177; ORF181; ORF183; ORF184; ORF186; ORF187; ORF188; ORF190; ORF191; ORF192; ORF194; ORF195; ORF196; ORF197; ORF198; ORF199; ORF201; ORF202; ORF204; ORF206; ORF207; ORF209; ORF212; ORF213; ORF217; ORF219;

ORF220; ORF221; ORF222; ORF223; ORF224; ORF225; ORF227;
ORF228; ORF231; ORF232; ORF234; ORF236; ORF237; ORF243;
ORF244; ORF245; ORF247; ORF248; ORF249; ORF252; ORF254;
ORF257; ORF260; ORF261; ORF263; ORF265; ORF266; ORF267;
ORF270; ORF271; ORF272; ORF274; ORF276; ORF277; ORF278;
ORF279; ORF282; ORF283; ORF284; ORF285; ORF287; ORF289;
ORF290; ORF291; ORF294; ORF298; ORF305; ORF306; ORF310;
ORF311; ORF313; ORF315; ORF316; ORF319; ORF320; ORF322;
ORF323; ORF325; ORF326; ORF327; ORF328; ORF330; ORF331;
ORF332; ORF333; ORF334; ORF335; ORF336; ORF338; ORF339;
ORF340; ORF341; ORF344; ORF345; ORF348; ORF349; ORF350;
ORF351; ORF352; ORF353; ORF356; ORF357; ORF358; ORF361;
ORF362; ORF366; ORF367; ORF368; ORF370; ORF372; ORF373;
ORF375; ORF377; ORF378; ORF379; ORF380; ORF382; ORF383;
ORF384; ORF385; ORF387; ORF389; ORF390; ORF391; ORF393;
ORF396; ORF398; ORF399; ORF403; ORF404; ORF406; ORF407;
ORF413; ORF414; ORF417; ORF418; ORF420; ORF421; ORF424;
ORF426; ORF427; ORF428; ORF430; ORF433; ORF434; ORF435;
ORF436; ORF437; ORF440; ORF443; ORF446; ORF448; ORF450;
ORF451; ORF454; ORF455; ORF457; ORF458; ORF459; ORF463;
ORF464; ORF466; ORF467; ORF468; ORF469; ORF470; ORF473;
ORF474; ORF475; ORF476; ORF477; ORF479; ORF480; ORF481;
ORF483; ORF484; ORF485; ORF486; ORF487; ORF488; ORF491;
ORF493; ORF496; ORF497; ORF498; ORF500; ORF501; ORF503;
ORF504; ORF508; ORF512; ORF513; ORF514; ORF519; ORF521;
ORF523; ORF524; ORF526; ORF527; ORF529; ORF530; ORF531;
ORF532; ORF534; ORF536; ORF537; ORF538; ORF540; ORF541;
ORF542; ORF543; ORF544; ORF545; ORF546; ORF547; ORF551;
ORF552; ORF553; ORF555; ORF558; ORF559; ORF560; ORF561;
ORF562; ORF566; ORF567; ORF568; ORF569; ORF571; ORF572;
ORF574; ORF575; ORF576; ORF580; ORF582; ORF585; ORF587;
ORF589; ORF592; ORF593; ORF595; ORF596; ORF597; ORF599;
ORF601; ORF602; ORF603; ORF604; ORF608; ORF609; ORF610;
ORF611; ORF615; ORF616; ORF617; ORF618; ORF621; ORF622;
ORF623; ORF624; ORF625; ORF628; ORF632; ORF633; ORF634;
ORF635; ORF637; ORF638; ORF640; ORF641; ORF643; ORF646;
ORF648; ORF649; ORF651; ORF652; ORF653; ORF654; ORF655;
ORF658; ORF664; ORF665; ORF666; ORF668; ORF669; ORF670;
ORF671; ORF672; ORF673; ORF674; ORF676; ORF677; ORF678;

ORF680; ORF682; ORF683; ORF684; ORF686; ORF688; ORF689;
ORF690; ORF691; ORF692; ORF693; ORF695; ORF696; ORF698;
ORF701; ORF703; ORF704; ORF705; ORF706; ORF707; ORF709;
ORF710; ORF711; ORF712; ORF713; ORF714; ORF715; ORF717;
ORF718; ORF720; ORF721; ORF722; ORF724; ORF726; ORF728;
ORF729; ORF730; ORF731; ORF732; ORF733; ORF734; ORF737;
ORF738; ORF739; ORF740; ORF742; ORF743; ORF744; ORF745;
ORF746; ORF748; ORF750; ORF751; ORF752; ORF753; ORF754;
ORF755; ORF757; ORF758; ORF759; ORF760; ORF764; ORF766;
ORF768; ORF769; ORF771; ORF772; ORF773; ORF774; ORF775;
ORF776; ORF777; ORF778; ORF779; ORF780; ORF781; ORF782;
ORF783; ORF786; ORF787; ORF788; ORF789; ORF790; ORF793;
ORF798; ORF800; ORF802; ORF803; ORF806; ORF808; ORF809;
ORF810; ORF811; ORF813; ORF814; ORF817; ORF820; ORF822;
ORF824; ORF825; ORF827; ORF828; ORF829; ORF830; ORF833;
ORF834; ORF835; ORF837; ORF838; ORF839; ORF840; ORF841;
ORF842; ORF843; ORF845; ORF848; ORF849; ORF850; ORF851;
ORF852; ORF854; ORF855; ORF856; ORF857; ORF859; ORF860;
ORF862; ORF863; ORF864; ORF866; ORF869; ORF872; ORF873;
ORF874; ORF878; ORF879; ORF880; ORF881; ORF883; ORF884;
ORF885; ORF886; ORF887; ORF892; ORF893; ORF894; ORF895;
ORF897; ORF899; ORF900; ORF901; ORF904; ORF906; ORF909;
ORF910; ORF912; ORF914; ORF917; ORF920; ORF921; ORF922;
ORF923; ORF924; ORF925; ORF926; ORF927; ORF930; ORF933;
ORF934; ORF935; ORF936; ORF937; ORF940; ORF941; ORF942;
ORF943; ORF944; ORF945; ORF947; ORF948; ORF951; ORF952;
ORF953; ORF954; ORF955; ORF956; ORF957; ORF958; ORF960;
ORF961; ORF962; ORF963; ORF964; ORF966; ORF967; ORF969;
ORF970; ORF971; ORF973; ORF974; ORF979; ORF980; ORF981;
ORF982; ORF984; ORF988; ORF989; ORF990; ORF991; ORF995;
ORF996; ORF999; ORF1001; ORF1003; ORF1004; ORF1005;
ORF1006; ORF1007; ORF1009; ORF1010; ORF1011; ORF1012;
ORF1013; ORF1014; ORF1016; ORF1017; ORF1018; ORF1020;
ORF1021; ORF1025; ORF1026; ORF1027; ORF1029; ORF1030;
ORF1031; ORF1035; ORF1036; ORF1037; ORF1038; ORF1039;
ORF1040; ORF1044; ORF1045; ORF1047; ORF1048; ORF1050;
ORF1051; ORF1052; ORF1053; ORF1055; ORF1056; ORF1057;
ORF1058; ORF1061; ORF1062; ORF1063; ORF1064; ORF1065;
ORF1066; ORF1068; ORF1069; ORF1072; ORF1074; ORF1076

and one of their representative fragments.

Preferably, the invention relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 4 and 6 transmembrane domains and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF7; ORF14; ORF16; ORF32; ORF34; ORF36; ORF38; ORF50; ORF57; ORF59; ORF61; ORF62; ORF63; ORF64; ORF67; ORF69; ORF72; ORF77; ORF80; ORF84; ORF87; ORF93; ORF95; ORF99; ORF108; ORF119; ORF125; ORF126; ORF129; ORF131; ORF136; ORF139; ORF146; ORF152; ORF154; ORF160; ORF161; ORF172; ORF179; ORF182; ORF185; ORF200; ORF203; ORF205; ORF239; ORF242; ORF250; ORF253; ORF256; ORF259; ORF262; ORF268; ORF275; ORF281; ORF286; ORF288; ORF292; ORF295; ORF296; ORF297; ORF299; ORF300; ORF308; ORF314; ORF317; ORF318; ORF324; ORF342; ORF343; ORF355; ORF360; ORF374; ORF376; ORF386; ORF388; ORF392; ORF394; ORF395; ORF402; ORF405; ORF411; ORF415; ORF416; ORF422; ORF423; ORF429; ORF432; ORF441; ORF442; ORF444; ORF449; ORF452; ORF456; ORF460; ORF461; ORF465; ORF471; ORF472; ORF482; ORF489; ORF492; ORF494; ORF495; ORF502; ORF505; ORF506; ORF509; ORF516; ORF517; ORF520; ORF525; ORF533; ORF539; ORF549; ORF554; ORF557; ORF563; ORF570; ORF573; ORF581; ORF590; ORF591; ORF600; ORF607; ORF612; ORF613; ORF620; ORF626; ORF629; ORF630; ORF639; ORF644; ORF647; ORF656; ORF659; ORF661; ORF685; ORF687; ORF699; ORF700; ORF708; ORF716; ORF719; ORF725; ORF747; ORF749; ORF756; ORF765; ORF767; ORF794; ORF796; ORF797; ORF799; ORF801; ORF807; ORF821; ORF823; ORF826; ORF847; ORF853; ORF861; ORF870; ORF871; ORF875; ORF882; ORF888; ORF889; ORF898; ORF902; ORF903; ORF911; ORF916; ORF931; ORF939; ORF975; ORF976; ORF978; ORF983; ORF986; ORF987; ORF992; ORF993; ORF1000; ORF1002; ORF1008; ORF1019; ORF1022; ORF1032; ORF1034; ORF1046; ORF1054; ORF1060; ORF1071 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention,

characterized in that they encode a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having at least 7 transmembrane domains and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF4; ORF6; ORF13; ORF20; ORF51; ORF71; ORF88; ORF118; ORF128; ORF132; ORF133; ORF158; ORF159; ORF174; ORF180; ORF189; ORF210; ORF211; ORF214; ORF215; ORF226; ORF229; ORF233; ORF235; ORF240; ORF246; ORF251; ORF255; ORF273; ORF354; ORF364; ORF369; ORF371; ORF397; ORF401; ORF409; ORF412; ORF419; ORF439; ORF453; ORF462; ORF490; ORF510; ORF511; ORF518; ORF535; ORF548; ORF550; ORF564; ORF565; ORF578; ORF579; ORF614; ORF631; ORF636; ORF650; ORF662; ORF667; ORF679; ORF681; ORF702; ORF727; ORF741; ORF763; ORF791; ORF792; ORF815; ORF816; ORF832; ORF846; ORF858; ORF865; ORF867; ORF868; ORF877; ORF891; ORF896; ORF907; ORF908; ORF918; ORF919; ORF932; ORF959; ORF977; ORF994; ORF998; ORF1024; ORF1028; ORF1042; ORF1067; ORF1070; ORF1073 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, such as for example triose phosphate isomerase or pyruvate kinase, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF10; ORF44; ORF45; ORF46; ORF47; ORF93; ORF101; ORF102; ORF103; ORF106; ORF107; ORF120; ORF121; ORF130; ORF135; ORF140; ORF143; ORF144; ORF145; ORF158; ORF159; ORF160; ORF161; ORF192; ORF193; ORF196; ORF196; ORF197; ORF198; ORF199; ORF227; ORF229; ORF236; ORF236; ORF239; ORF243; ORF245; ORF264; ORF265; ORF297; ORF331; ORF333; ORF359; ORF360; ORF374; ORF404; ORF405; ORF405; ORF410; ORF415; ORF415; ORF416; ORF417; ORF432; ORF460; ORF461; ORF462; ORF495; ORF513; ORF515; ORF566; ORF566; ORF566; ORF589; ORF613; ORF645; ORF646; ORF647; ORF652; ORF653;

ORF654; ORF672; ORF673; ORF674; ORF682; ORF684; ORF692; ORF700; ORF725; ORF801; ORF802; ORF835; ORF836; ORF837; ORF860; ORF861; ORF862; ORF863; ORF869; ORF869; ORF925; ORF964; ORF983 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleotides, such as for example CTP synthetase or GMP synthetase, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF142; ORF142; ORF169; ORF256; ORF268; ORF325; ORF352; ORF366; ORF435; ORF444; ORF528; ORF529; ORF530; ORF548; ORF549; ORF601; ORF602; ORF617; ORF619; ORF644; ORF745; ORF971; ORF972; ORF1023 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleic acids, such as for example DNA polymerases or DNA topoisomerases, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF5; ORF12; ORF82; ORF96; ORF97; ORF98; ORF99; ORF100; ORF105; ORF118; ORF136; ORF137; ORF163; ORF190; ORF204; ORF259; ORF260; ORF262; ORF290; ORF300; ORF301; ORF302; ORF387; ORF427; ORF434; ORF441; ORF444; ORF471; ORF595; ORF596; ORF597; ORF599; ORF600; ORF605; ORF612; ORF624; ORF625; ORF650; ORF657; ORF658; ORF702; ORF703; ORF704; ORF708; ORF719; ORF766; ORF767; ORF775; ORF779; ORF787; ORF788; ORF794; ORF841; ORF842; ORF883; ORF884; ORF907; ORF918; ORF924; ORF928; ORF929; ORF962; ORF962; ORF963; ORF969; ORF970; ORF975; ORF979; ORF995; ORF1031; ORF1032 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention,

characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of amino acids, such as for example serine hydroxymethyl transferase or the proteins which load amino acids onto transfer RNAs, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF27; ORF41; ORF55; ORF56; ORF57; ORF59; ORF62; ORF63; ORF64; ORF65; ORF119; ORF132; ORF240; ORF241; ORF277; ORF278; ORF279; ORF382; ORF406; ORF428; ORF442; ORF446; ORF447; ORF453; ORF454; ORF541; ORF542; ORF591; ORF608; ORF609; ORF610; ORF618; ORF648; ORF649; ORF660; ORF661; ORF677; ORF717; ORF765; ORF797; ORF871; ORF875; ORF920; ORF922; ORF937; ORF998; ORF1020; ORF1021; ORF1034; ORF1044; ORF1046; ORF1049 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of polypeptides, such as for example protein kinases or proteases, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF21; ORF21; ORF22; ORF23; ORF24; ORF25; ORF26; ORF75; ORF84; ORF84; ORF86; ORF92; ORF133; ORF151; ORF152; ORF157; ORF179; ORF209; ORF307; ORF326; ORF343; ORF344; ORF345; ORF371; ORF429; ORF519; ORF557; ORF586; ORF587; ORF630; ORF656; ORF706; ORF707; ORF730; ORF751; ORF752; ORF786; ORF847; ORF885; ORF923; ORF978; ORF1039; ORF1048 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of fatty acids, such as for example succinyl-CoA-synthesizing proteins or phosphatidylserine synthetase, and in that they comprise a nucleotide sequence chosen from the

following sequences:
ORF4; ORF15; ORF16; ORF141; ORF173; ORF205; ORF205; ORF206; ORF207; ORF208; ORF312; ORF355; ORF415; ORF550; ORF558; ORF560; ORF561; ORF574; ORF574; ORF577; ORF578; ORF590; ORF614; ORF772; ORF808; ORF809; ORF904; ORF905; ORF905; ORF933; ORF934; ORF934; ORF936 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the synthesis of the wall, such as for example KDO transferase, and the proteins responsible for the attachment of certain sugars onto the exposed proteins, and in that they comprise a nucleotide sequence chosen from the following sequences:
ORF87; ORF196; ORF242; ORF269; ORF628; ORF629; ORF634; ORF635; ORF637; ORF638; ORF1019 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the transcription, translation and/or maturation process, such as for example initiation factors, RNA polymerases or certain chaperone proteins, and in that they comprise a nucleotide sequence chosen from the following sequences:
ORF112; ORF113; ORF332; ORF212; ORF213; ORF350; ORF362; ORF363; ORF364; ORF407; ORF451; ORF546; ORF643; ORF744; ORF746; ORF833; ORF868; ORF981; ORF982; ORF1003; ORF1011; ORF1042 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* ribosomal polypeptide or one of its fragments, such as for example the ribosomal proteins L21, L27 and S10, and in that they comprise a

nucleotide sequence chosen from the following sequences:

ORF114; ORF115; ORF116; ORF328; ORF361; ORF375; ORF445; ORF543; ORF584; ORF585; ORF743; ORF813; ORF941; ORF942; ORF944; ORF946; ORF947; ORF948; ORF950; ORF951; ORF952; ORF953; ORF954; ORF955; ORF955; ORF957; ORF958; ORF960; ORF961; ORF1040; ORF1041; ORF1043; ORF1063; ORF1064 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* transport polypeptide or one of its fragments, such as for example the proteins for transporting amino acids, sugars and certain oligopeptides, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF6; ORF50; ORF51; ORF80; ORF125; ORF126; ORF128; ORF129; ORF215; ORF246; ORF248; ORF249; ORF251; ORF252; ORF253; ORF255; ORF271; ORF275; ORF293; ORF309; ORF323; ORF324; ORF398; ORF401; ORF449; ORF511; ORF512; ORF564; ORF565; ORF667; ORF679; ORF680; ORF711; ORF712; ORF713; ORF714; ORF715; ORF730; ORF731; ORF736; ORF737; ORF738; ORF870; ORF908; ORF919; ORF977; ORF987; ORF988; ORF992; ORF993; ORF994; ORF1028; ORF1029 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the virulence process, such as for example the proteins analogous to the *Escherichia coli* vacB protein, and in that they comprise a nucleotide sequence chosen from the following sequences:

ORF20; ORF815; ORF816; ORF898; ORF1059; ORF1060 and one of their representative fragments.

Preferably, the invention also relates to the nucleotide sequences according to the invention, characterized in that they encode a *Chlamydia trachomatis* polypeptide or one of its fragments which

is involved in the secretory system and/or which is secreted, such as for example proteins homologous to proteins in the secretory system of certain bacteria such as the Salmonellae or the Yersiniae, and in that they comprise a nucleotide sequence chosen from the following sequences:
ORF758; ORF888; ORF889; ORF890; ORF891; ORF896; ORF897; ORF898 and one of their representative fragments.

Preferably, the invention also relates to nucleotide sequences according to the invention, characterized in that they encode a polypeptide specific to *Chlamydiae* or one of its fragments, and in that they comprise a nucleotide sequence chosen from the following sequences:
ORF22; ORF29; ORF31; ORF32; ORF34; ORF35; ORF39; ORF40; ORF43; ORF48; ORF49; ORF50; ORF52; ORF53; ORF54; ORF72; ORF77; ORF78; ORF87; ORF90; ORF95; ORF108; ORF110; ORF111; ORF122; ORF123; ORF124; ORF127; ORF138; ORF144; ORF146; ORF153; ORF155; ORF164; ORF166; ORF175; ORF182; ORF184; ORF186; ORF187; ORF188; ORF202; ORF210; ORF247; ORF258; ORF266; ORF267; ORF270; ORF273; ORF274; ORF295; ORF296; ORF305; ORF306; ORF309; ORF318; ORF319; ORF322; ORF326; ORF342; ORF357; ORF376; ORF379; ORF380; ORF388; ORF390; ORF400; ORF431; ORF433; ORF438; ORF443; ORF456; ORF457; ORF458; ORF464; ORF468; ORF470; ORF473; ORF486; ORF489; ORF497; ORF501; ORF503; ORF504; ORF508; ORF512; ORF521; ORF522; ORF523; ORF524; ORF533; ORF535; ORF536; ORF537; ORF538; ORF539; ORF540; ORF554; ORF563; ORF572; ORF579; ORF595; ORF603; ORF604; ORF606; ORF607; ORF615; ORF616; ORF622; ORF641; ORF642; ORF659; ORF668; ORF670; ORF693; ORF695; ORF696; ORF699; ORF703; ORF704; ORF716; ORF726; ORF728; ORF739; ORF742; ORF747; ORF750; ORF751; ORF755; ORF757; ORF759; ORF761; ORF762; ORF763; ORF764; ORF773; ORF780; ORF781; ORF789; ORF800; ORF803; ORF804; ORF818; ORF820; ORF822; ORF823; ORF824; ORF827; ORF828; ORF839; ORF849; ORF850; ORF851; ORF852; ORF855; ORF856; ORF857; ORF858; ORF859; ORF860; ORF861; ORF862; ORF863; ORF865; ORF868; ORF869; ORF870; ORF871; ORF872; ORF873; ORF874; ORF875; ORF877; ORF878; ORF880; ORF882; ORF884;

ORF886; ORF893; ORF901; ORF906; ORF910; ORF912; ORF915; ORF916; ORF917; ORF926; ORF929; ORF933; ORF965; ORF967; ORF968; ORF984; ORF986; ORF989; ORF990; ORF996; ORF997; ORF1001; ORF1002; ORF1013; ORF1016; ORF1031; ORF1033; ORF1035; ORF1049; ORF1051; ORF1052; ORF1054; ORF1056; ORF1057; ORF1058; ORF1062; ORF1070; ORF1071; ORF1073 and one of their representative fragments.

The subject of the invention is also a polypeptide according to the invention, characterized in that it is a polypeptide of the cellular envelope, preferably of the outer cellular envelope, of *Chlamydia trachomatis* or one of its fragments. According to the invention, the said polypeptide is preferably chosen from the polypeptides having the following sequences:

SEQ ID No. 3; SEQ ID No. 19; SEQ ID No. 51; SEQ ID No. 189; SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 324; SEQ ID No. 477; SEQ ID No. 478; SEQ ID No. 479; SEQ ID No. 481; SEQ ID No. 482; SEQ ID No. 483; SEQ ID No. 484; SEQ ID No. 486; SEQ ID No. 488; SEQ ID No. 489; SEQ ID No. 490; SEQ ID No. 572; SEQ ID No. 573; SEQ ID No. 742; SEQ ID No. 817; SEQ ID No. 818; SEQ ID No. 820; SEQ ID No. 1035; SEQ ID No. 1036; SEQ ID No. 1037; SEQ ID No. 1038; SEQ ID No. 1070; SEQ ID No. 1071; SEQ ID No. 1073 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 1 and 3 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 2; SEQ ID No. 3; SEQ ID No. 5; SEQ ID No. 8; SEQ ID No. 9; SEQ ID No. 10; SEQ ID No. 11; SEQ ID No. 12; SEQ ID No. 17; SEQ ID No. 21; SEQ ID No. 26; SEQ ID No. 27; SEQ ID No. 28; SEQ ID No. 29; SEQ ID No. 30; SEQ ID No. 31; SEQ ID No. 33; SEQ ID No. 35; SEQ ID No. 37; SEQ ID No. 39; SEQ ID No. 40; SEQ ID No. 41; SEQ ID No. 42; SEQ ID No. 43; SEQ ID No. 44;

SEQ ID No. 45; SEQ ID No. 46; SEQ ID No. 47;
SEQ ID No. 48; SEQ ID No. 49; SEQ ID No. 52;
SEQ ID No. 53; SEQ ID No. 55; SEQ ID No. 56;
SEQ ID No. 58; SEQ ID No. 65; SEQ ID No. 66;
SEQ ID No. 68; SEQ ID No. 70; SEQ ID No. 74;
SEQ ID No. 75; SEQ ID No. 76; SEQ ID No. 78;
SEQ ID No. 79; SEQ ID No. 81; SEQ ID No. 82;
SEQ ID No. 83; SEQ ID No. 86; SEQ ID No. 91;
SEQ ID No. 92; SEQ ID No. 94; SEQ ID No. 97;
SEQ ID No. 100; SEQ ID No. 102; SEQ ID No. 103;
SEQ ID No. 105; SEQ ID No. 106; SEQ ID No. 107;
SEQ ID No. 109; SEQ ID No. 110; SEQ ID No. 111;
SEQ ID No. 112; SEQ ID No. 113; SEQ ID No. 114;
SEQ ID No. 115; SEQ ID No. 116; SEQ ID No. 117;
SEQ ID No. 120; SEQ ID No. 122; SEQ ID No. 123;
SEQ ID No. 130; SEQ ID No. 134; SEQ ID No. 135;
SEQ ID No. 137; SEQ ID No. 140; SEQ ID No. 141;
SEQ ID No. 143; SEQ ID No. 144; SEQ ID No. 145;
SEQ ID No. 147; SEQ ID No. 148; SEQ ID No. 149;
SEQ ID No. 150; SEQ ID No. 151; SEQ ID No. 155;
SEQ ID No. 156; SEQ ID No. 162; SEQ ID No. 163;
SEQ ID No. 164; SEQ ID No. 165; SEQ ID No. 166;
SEQ ID No. 167; SEQ ID No. 168; SEQ ID No. 169;
SEQ ID No. 170; SEQ ID No. 171; SEQ ID No. 173;
SEQ ID No. 175; SEQ ID No. 176; SEQ ID No. 177;
SEQ ID No. 181; SEQ ID No. 183; SEQ ID No. 184;
SEQ ID No. 186; SEQ ID No. 187; SEQ ID No. 188;
SEQ ID No. 190; SEQ ID No. 191; SEQ ID No. 192;
SEQ ID No. 194; SEQ ID No. 195; SEQ ID No. 196;
SEQ ID No. 197; SEQ ID No. 198; SEQ ID No. 199;
SEQ ID No. 201; SEQ ID No. 202; SEQ ID No. 204;
SEQ ID No. 206; SEQ ID No. 207; SEQ ID No. 209;
SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 217;
SEQ ID No. 219; SEQ ID No. 220; SEQ ID No. 221;
SEQ ID No. 222; SEQ ID No. 223; SEQ ID No. 224;
SEQ ID No. 225; SEQ ID No. 227; SEQ ID No. 228;
SEQ ID No. 231; SEQ ID No. 232; SEQ ID No. 234;
SEQ ID No. 236; SEQ ID No. 237; SEQ ID No. 243;
SEQ ID No. 244; SEQ ID No. 245; SEQ ID No. 247;

SEQ ID No. 248; SEQ ID No. 249; SEQ ID No. 252;
SEQ ID No. 254; SEQ ID No. 257; SEQ ID No. 260;
SEQ ID No. 261; SEQ ID No. 263; SEQ ID No. 265;
SEQ ID No. 266; SEQ ID No. 267; SEQ ID No. 270;
SEQ ID No. 271; SEQ ID No. 272; SEQ ID No. 274;
SEQ ID No. 276; SEQ ID No. 277; SEQ ID No. 278;
SEQ ID No. 279; SEQ ID No. 282; SEQ ID No. 283;
SEQ ID No. 284; SEQ ID No. 285; SEQ ID No. 287;
SEQ ID No. 289; SEQ ID No. 290; SEQ ID No. 291;
SEQ ID No. 294; SEQ ID No. 298; SEQ ID No. 305;
SEQ ID No. 306; SEQ ID No. 310; SEQ ID No. 311;
SEQ ID No. 313; SEQ ID No. 315; SEQ ID No. 316;
SEQ ID No. 319; SEQ ID No. 320; SEQ ID No. 322;
SEQ ID No. 323; SEQ ID No. 325; SEQ ID No. 326;
SEQ ID No. 327; SEQ ID No. 328; SEQ ID No. 330;
SEQ ID No. 331; SEQ ID No. 332; SEQ ID No. 333;
SEQ ID No. 334; SEQ ID No. 335; SEQ ID No. 336;
SEQ ID No. 338; SEQ ID No. 339; SEQ ID No. 340;
SEQ ID No. 341; SEQ ID No. 344; SEQ ID No. 345;
SEQ ID No. 348; SEQ ID No. 349; SEQ ID No. 350;
SEQ ID No. 351; SEQ ID No. 352; SEQ ID No. 353;
SEQ ID No. 356; SEQ ID No. 357; SEQ ID No. 358;
SEQ ID No. 361; SEQ ID No. 362; SEQ ID No. 366;
SEQ ID No. 367; SEQ ID No. 368; SEQ ID No. 370;
SEQ ID No. 372; SEQ ID No. 373; SEQ ID No. 375;
SEQ ID No. 377; SEQ ID No. 378; SEQ ID No. 379;
SEQ ID No. 380; SEQ ID No. 382; SEQ ID No. 383;
SEQ ID No. 384; SEQ ID No. 385; SEQ ID No. 387;
SEQ ID No. 389; SEQ ID No. 390; SEQ ID No. 391;
SEQ ID No. 393; SEQ ID No. 396; SEQ ID No. 398;
SEQ ID No. 399; SEQ ID No. 403; SEQ ID No. 404;
SEQ ID No. 406; SEQ ID No. 407; SEQ ID No. 413;
SEQ ID No. 414; SEQ ID No. 417; SEQ ID No. 418;
SEQ ID No. 420; SEQ ID No. 421; SEQ ID No. 424;
SEQ ID No. 426; SEQ ID No. 427; SEQ ID No. 428;
SEQ ID No. 430; SEQ ID No. 433; SEQ ID No. 434;
SEQ ID No. 435; SEQ ID No. 436; SEQ ID No. 437;
SEQ ID No. 440; SEQ ID No. 443; SEQ ID No. 446;
SEQ ID No. 448; SEQ ID No. 450; SEQ ID No. 451;

SEQ ID No. 454; SEQ ID No. 455; SEQ ID No. 457;
SEQ ID No. 458; SEQ ID No. 459; SEQ ID No. 463;
SEQ ID No. 464; SEQ ID No. 466; SEQ ID No. 467;
SEQ ID No. 468; SEQ ID No. 469; SEQ ID No. 470;
SEQ ID No. 473; SEQ ID No. 474; SEQ ID No. 475;
SEQ ID No. 476; SEQ ID No. 477; SEQ ID No. 479;
SEQ ID No. 480; SEQ ID No. 481; SEQ ID No. 483;
SEQ ID No. 484; SEQ ID No. 485; SEQ ID No. 486;
SEQ ID No. 487; SEQ ID No. 488; SEQ ID No. 491;
SEQ ID No. 493; SEQ ID No. 496; SEQ ID No. 497;
SEQ ID No. 498; SEQ ID No. 500; SEQ ID No. 501;
SEQ ID No. 503; SEQ ID No. 504; SEQ ID No. 508;
SEQ ID No. 512; SEQ ID No. 513; SEQ ID No. 514;
SEQ ID No. 519; SEQ ID No. 521; SEQ ID No. 523;
SEQ ID No. 524; SEQ ID No. 526; SEQ ID No. 527;
SEQ ID No. 529; SEQ ID No. 530; SEQ ID No. 531;
SEQ ID No. 532; SEQ ID No. 534; SEQ ID No. 536;
SEQ ID No. 537; SEQ ID No. 538; SEQ ID No. 540;
SEQ ID No. 541; SEQ ID No. 542; SEQ ID No. 543;
SEQ ID No. 544; SEQ ID No. 545; SEQ ID No. 546;
SEQ ID No. 547; SEQ ID No. 551; SEQ ID No. 552;
SEQ ID No. 553; SEQ ID No. 555; SEQ ID No. 558;
SEQ ID No. 559; SEQ ID No. 560; SEQ ID No. 561;
SEQ ID No. 562; SEQ ID No. 566; SEQ ID No. 567;
SEQ ID No. 568; SEQ ID No. 569; SEQ ID No. 571;
SEQ ID No. 572; SEQ ID No. 574; SEQ ID No. 575;
SEQ ID No. 576; SEQ ID No. 580; SEQ ID No. 582;
SEQ ID No. 585; SEQ ID No. 587; SEQ ID No. 589;
SEQ ID No. 592; SEQ ID No. 593; SEQ ID No. 595;
SEQ ID No. 596; SEQ ID No. 597; SEQ ID No. 599;
SEQ ID No. 601; SEQ ID No. 602; SEQ ID No. 603;
SEQ ID No. 604; SEQ ID No. 608; SEQ ID No. 609;
SEQ ID No. 610; SEQ ID No. 611; SEQ ID No. 615;
SEQ ID No. 616; SEQ ID No. 617; SEQ ID No. 618;
SEQ ID No. 621; SEQ ID No. 622; SEQ ID No. 623;
SEQ ID No. 624; SEQ ID No. 625; SEQ ID No. 628;
SEQ ID No. 632; SEQ ID No. 633; SEQ ID No. 634;
SEQ ID No. 635; SEQ ID No. 637; SEQ ID No. 638;
SEQ ID No. 640; SEQ ID No. 641; SEQ ID No. 643;

SEQ ID No. 646; SEQ ID No. 648; SEQ ID No. 649;
SEQ ID No. 651; SEQ ID No. 652; SEQ ID No. 653;
SEQ ID No. 654; SEQ ID No. 655; SEQ ID No. 658;
SEQ ID No. 664; SEQ ID No. 665; SEQ ID No. 666;
SEQ ID No. 668; SEQ ID No. 669; SEQ ID No. 670;
SEQ ID No. 671; SEQ ID No. 672; SEQ ID No. 673;
SEQ ID No. 674; SEQ ID No. 676; SEQ ID No. 677;
SEQ ID No. 678; SEQ ID No. 680; SEQ ID No. 682;
SEQ ID No. 683; SEQ ID No. 684; SEQ ID No. 686;
SEQ ID No. 688; SEQ ID No. 689; SEQ ID No. 690;
SEQ ID No. 691; SEQ ID No. 692; SEQ ID No. 693;
SEQ ID No. 695; SEQ ID No. 696; SEQ ID No. 698;
SEQ ID No. 701; SEQ ID No. 703; SEQ ID No. 704;
SEQ ID No. 705; SEQ ID No. 706; SEQ ID No. 707;
SEQ ID No. 709; SEQ ID No. 710; SEQ ID No. 711;
SEQ ID No. 712; SEQ ID No. 713; SEQ ID No. 714;
SEQ ID No. 715; SEQ ID No. 717; SEQ ID No. 718;
SEQ ID No. 720; SEQ ID No. 721; SEQ ID No. 722;
SEQ ID No. 724; SEQ ID No. 726; SEQ ID No. 728;
SEQ ID No. 729; SEQ ID No. 730; SEQ ID No. 731;
SEQ ID No. 732; SEQ ID No. 733; SEQ ID No. 734;
SEQ ID No. 737; SEQ ID No. 738; SEQ ID No. 739;
SEQ ID No. 740; SEQ ID No. 742; SEQ ID No. 743;
SEQ ID No. 744; SEQ ID No. 745; SEQ ID No. 746;
SEQ ID No. 748; SEQ ID No. 750; SEQ ID No. 751;
SEQ ID No. 752; SEQ ID No. 753; SEQ ID No. 754;
SEQ ID No. 755; SEQ ID No. 757; SEQ ID No. 758;
SEQ ID No. 759; SEQ ID No. 760; SEQ ID No. 764;
SEQ ID No. 766; SEQ ID No. 768; SEQ ID No. 769;
SEQ ID No. 771; SEQ ID No. 772; SEQ ID No. 773;
SEQ ID No. 774; SEQ ID No. 775; SEQ ID No. 776;
SEQ ID No. 777; SEQ ID No. 778; SEQ ID No. 779;
SEQ ID No. 780; SEQ ID No. 781; SEQ ID No. 782;
SEQ ID No. 783; SEQ ID No. 786; SEQ ID No. 787;
SEQ ID No. 788; SEQ ID No. 789; SEQ ID No. 790;
SEQ ID No. 793; SEQ ID No. 798; SEQ ID No. 800;
SEQ ID No. 802; SEQ ID No. 803; SEQ ID No. 806;
SEQ ID No. 808; SEQ ID No. 809; SEQ ID No. 810;
SEQ ID No. 811; SEQ ID No. 813; SEQ ID No. 814;

SEQ ID No. 817; SEQ ID No. 820; SEQ ID No. 822;
SEQ ID No. 824; SEQ ID No. 825; SEQ ID No. 827;
SEQ ID No. 828; SEQ ID No. 829; SEQ ID No. 830;
SEQ ID No. 833; SEQ ID No. 834; SEQ ID No. 835;
SEQ ID No. 837; SEQ ID No. 838; SEQ ID No. 839;
SEQ ID No. 840; SEQ ID No. 841; SEQ ID No. 842;
SEQ ID No. 843; SEQ ID No. 845; SEQ ID No. 848;
SEQ ID No. 849; SEQ ID No. 850; SEQ ID No. 851;
SEQ ID No. 852; SEQ ID No. 854; SEQ ID No. 855;
SEQ ID No. 856; SEQ ID No. 857; SEQ ID No. 859;
SEQ ID No. 860; SEQ ID No. 862; SEQ ID No. 863;
SEQ ID No. 864; SEQ ID No. 866; SEQ ID No. 869;
SEQ ID No. 872; SEQ ID No. 873; SEQ ID No. 874;
SEQ ID No. 878; SEQ ID No. 879; SEQ ID No. 880;
SEQ ID No. 881; SEQ ID No. 883; SEQ ID No. 884;
SEQ ID No. 885; SEQ ID No. 886; SEQ ID No. 887;
SEQ ID No. 892; SEQ ID No. 893; SEQ ID No. 894;
SEQ ID No. 895; SEQ ID No. 897; SEQ ID No. 899;
SEQ ID No. 900; SEQ ID No. 901; SEQ ID No. 904;
SEQ ID No. 906; SEQ ID No. 909; SEQ ID No. 910;
SEQ ID No. 912; SEQ ID No. 914; SEQ ID No. 917;
SEQ ID No. 920; SEQ ID No. 921; SEQ ID No. 922;
SEQ ID No. 923; SEQ ID No. 924; SEQ ID No. 925;
SEQ ID No. 926; SEQ ID No. 927; SEQ ID No. 930;
SEQ ID No. 933; SEQ ID No. 934; SEQ ID No. 935;
SEQ ID No. 936; SEQ ID No. 937; SEQ ID No. 940;
SEQ ID No. 941; SEQ ID No. 942; SEQ ID No. 943;
SEQ ID No. 944; SEQ ID No. 945; SEQ ID No. 947;
SEQ ID No. 948; SEQ ID No. 951; SEQ ID No. 952;
SEQ ID No. 953; SEQ ID No. 954; SEQ ID No. 955;
SEQ ID No. 956; SEQ ID No. 957; SEQ ID No. 958;
SEQ ID No. 960; SEQ ID No. 961; SEQ ID No. 962;
SEQ ID No. 963; SEQ ID No. 964; SEQ ID No. 966;
SEQ ID No. 967; SEQ ID No. 969; SEQ ID No. 970;
SEQ ID No. 971; SEQ ID No. 973; SEQ ID No. 974;
SEQ ID No. 979; SEQ ID No. 980; SEQ ID No. 981;
SEQ ID No. 982; SEQ ID No. 984; SEQ ID No. 988;
SEQ ID No. 989; SEQ ID No. 990; SEQ ID No. 991;
SEQ ID No. 995; SEQ ID No. 996; SEQ ID No. 999;

SEQ ID No. 1001; SEQ ID No. 1003; SEQ ID No. 1004;
SEQ ID No. 1005; SEQ ID No. 1006; SEQ ID No. 1007;
SEQ ID No. 1009; SEQ ID No. 1010; SEQ ID No. 1011;
SEQ ID No. 1012; SEQ ID No. 1013; SEQ ID No. 1014;
SEQ ID No. 1016; SEQ ID No. 1017; SEQ ID No. 1018;
SEQ ID No. 1020; SEQ ID No. 1021; SEQ ID No. 1025;
SEQ ID No. 1026; SEQ ID No. 1027; SEQ ID No. 1029;
SEQ ID No. 1030; SEQ ID No. 1031; SEQ ID No. 1035;
SEQ ID No. 1036; SEQ ID No. 1037; SEQ ID No. 1038;
SEQ ID No. 1039; SEQ ID No. 1040; SEQ ID No. 1044;
SEQ ID No. 1045; SEQ ID No. 1047; SEQ ID No. 1048;
SEQ ID No. 1050; SEQ ID No. 1051; SEQ ID No. 1052;
SEQ ID No. 1053; SEQ ID No. 1055; SEQ ID No. 1056;
SEQ ID No. 1057; SEQ ID No. 1058; SEQ ID No. 1061;
SEQ ID No. 1062; SEQ ID No. 1063; SEQ ID No. 1064;
SEQ ID No. 1065; SEQ ID No. 1066; SEQ ID No. 1068;
SEQ ID No. 1069; SEQ ID No. 1072; SEQ ID No. 1074;
SEQ ID No. 1076 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 4 and 6 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 7; SEQ ID No. 14; SEQ ID No. 16;
SEQ ID No. 32; SEQ ID No. 34; SEQ ID No. 36;
SEQ ID No. 38; SEQ ID No. 50; SEQ ID No. 57;
SEQ ID No. 59; SEQ ID No. 61; SEQ ID No. 62;
SEQ ID No. 63; SEQ ID No. 64; SEQ ID No. 67;
SEQ ID No. 69; SEQ ID No. 72; SEQ ID No. 77;
SEQ ID No. 80; SEQ ID No. 84; SEQ ID No. 87;
SEQ ID No. 93; SEQ ID No. 95; SEQ ID No. 99;
SEQ ID No. 108; SEQ ID No. 119; SEQ ID No. 125;
SEQ ID No. 126; SEQ ID No. 129; SEQ ID No. 131;
SEQ ID No. 136; SEQ ID No. 139; SEQ ID No. 146;
SEQ ID No. 152; SEQ ID No. 154; SEQ ID No. 160;
SEQ ID No. 161; SEQ ID No. 172; SEQ ID No. 179;
SEQ ID No. 182; SEQ ID No. 185; SEQ ID No. 200;
SEQ ID No. 203; SEQ ID No. 205; SEQ ID No. 239;

SEQ ID No. 242; SEQ ID No. 250; SEQ ID No. 253;
SEQ ID No. 256; SEQ ID No. 259; SEQ ID No. 262;
SEQ ID No. 268; SEQ ID No. 275; SEQ ID No. 281;
SEQ ID No. 286; SEQ ID No. 288; SEQ ID No. 292;
SEQ ID No. 295; SEQ ID No. 296; SEQ ID No. 297;
SEQ ID No. 299; SEQ ID No. 300; SEQ ID No. 308;
SEQ ID No. 314; SEQ ID No. 317; SEQ ID No. 318;
SEQ ID No. 324; SEQ ID No. 342; SEQ ID No. 343;
SEQ ID No. 355; SEQ ID No. 360; SEQ ID No. 374;
SEQ ID No. 376; SEQ ID No. 386; SEQ ID No. 388;
SEQ ID No. 392; SEQ ID No. 394; SEQ ID No. 395;
SEQ ID No. 402; SEQ ID No. 405; SEQ ID No. 411;
SEQ ID No. 415; SEQ ID No. 416; SEQ ID No. 422;
SEQ ID No. 423; SEQ ID No. 429; SEQ ID No. 432;
SEQ ID No. 441; SEQ ID No. 442; SEQ ID No. 444;
SEQ ID No. 449; SEQ ID No. 452; SEQ ID No. 456;
SEQ ID No. 460; SEQ ID No. 461; SEQ ID No. 465;
SEQ ID No. 471; SEQ ID No. 472; SEQ ID No. 482;
SEQ ID No. 489; SEQ ID No. 492; SEQ ID No. 494;
SEQ ID No. 495; SEQ ID No. 502; SEQ ID No. 505;
SEQ ID No. 506; SEQ ID No. 509; SEQ ID No. 516;
SEQ ID No. 517; SEQ ID No. 520; SEQ ID No. 525;
SEQ ID No. 533; SEQ ID No. 539; SEQ ID No. 549;
SEQ ID No. 554; SEQ ID No. 557; SEQ ID No. 563;
SEQ ID No. 570; SEQ ID No. 573; SEQ ID No. 581;
SEQ ID No. 590; SEQ ID No. 591; SEQ ID No. 600;
SEQ ID No. 607; SEQ ID No. 612; SEQ ID No. 613;
SEQ ID No. 620; SEQ ID No. 626; SEQ ID No. 629;
SEQ ID No. 630; SEQ ID No. 639; SEQ ID No. 644;
SEQ ID No. 647; SEQ ID No. 656; SEQ ID No. 659;
SEQ ID No. 661; SEQ ID No. 685; SEQ ID No. 687;
SEQ ID No. 699; SEQ ID No. 700; SEQ ID No. 708;
SEQ ID No. 716; SEQ ID No. 719; SEQ ID No. 725;
SEQ ID No. 747; SEQ ID No. 749; SEQ ID No. 756;
SEQ ID No. 765; SEQ ID No. 767; SEQ ID No. 794;
SEQ ID No. 796; SEQ ID No. 797; SEQ ID No. 799;
SEQ ID No. 801; SEQ ID No. 807; SEQ ID No. 821;
SEQ ID No. 823; SEQ ID No. 826; SEQ ID No. 847;
SEQ ID No. 853; SEQ ID No. 861; SEQ ID No. 870;

SEQ ID No. 871; SEQ ID No. 875; SEQ ID No. 882;
SEQ ID No. 888; SEQ ID No. 889; SEQ ID No. 898;
SEQ ID No. 902; SEQ ID No. 903; SEQ ID No. 911;
SEQ ID No. 916; SEQ ID No. 931; SEQ ID No. 939;
SEQ ID No. 975; SEQ ID No. 976; SEQ ID No. 978;
SEQ ID No. 983; SEQ ID No. 986; SEQ ID No. 987;
SEQ ID No. 992; SEQ ID No. 993; SEQ ID No. 1000;
SEQ ID No. 1002; SEQ ID No. 1008; SEQ ID No. 1019;
SEQ ID No. 1022; SEQ ID No. 1032; SEQ ID No. 1034;
SEQ ID No. 1046; SEQ ID No. 1054; SEQ ID No. 1060;
SEQ ID No. 1071 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having at least 7 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 4; SEQ ID No. 6; SEQ ID No. 13;
SEQ ID No. 20; SEQ ID No. 51; SEQ ID No. 71;
SEQ ID No. 88; SEQ ID No. 118; SEQ ID No. 128;
SEQ ID No. 132; SEQ ID No. 133; SEQ ID No. 158;
SEQ ID No. 159; SEQ ID No. 174; SEQ ID No. 180;
SEQ ID No. 189; SEQ ID No. 210; SEQ ID No. 211;
SEQ ID No. 214; SEQ ID No. 215; SEQ ID No. 226;
SEQ ID No. 229; SEQ ID No. 233; SEQ ID No. 235;
SEQ ID No. 240; SEQ ID No. 246; SEQ ID No. 251;
SEQ ID No. 255; SEQ ID No. 273; SEQ ID No. 354;
SEQ ID No. 364; SEQ ID No. 369; SEQ ID No. 371;
SEQ ID No. 397; SEQ ID No. 401; SEQ ID No. 409;
SEQ ID No. 412; SEQ ID No. 419; SEQ ID No. 439;
SEQ ID No. 453; SEQ ID No. 462; SEQ ID No. 490;
SEQ ID No. 510; SEQ ID No. 511; SEQ ID No. 518;
SEQ ID No. 535; SEQ ID No. 548; SEQ ID No. 550;
SEQ ID No. 564; SEQ ID No. 565; SEQ ID No. 578;
SEQ ID No. 579; SEQ ID No. 614; SEQ ID No. 631;
SEQ ID No. 636; SEQ ID No. 650; SEQ ID No. 662;
SEQ ID No. 667; SEQ ID No. 679; SEQ ID No. 681;
SEQ ID No. 702; SEQ ID No. 727; SEQ ID No. 741;
SEQ ID No. 763; SEQ ID No. 791; SEQ ID No. 792;

SEQ ID No. 815; SEQ ID No. 816; SEQ ID No. 832;
SEQ ID No. 846; SEQ ID No. 858; SEQ ID No. 865;
SEQ ID No. 867; SEQ ID No. 868; SEQ ID No. 877;
SEQ ID No. 891; SEQ ID No. 896; SEQ ID No. 907;
SEQ ID No. 908; SEQ ID No. 918; SEQ ID No. 919;
SEQ ID No. 932; SEQ ID No. 959; SEQ ID No. 977;
SEQ ID No. 994; SEQ ID No. 998; SEQ ID No. 1024;
SEQ ID No. 1028; SEQ ID No. 1042; SEQ ID No. 1067;
SEQ ID No. 1070; SEQ ID No. 1073 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 10; SEQ ID No. 44; SEQ ID No. 45;
SEQ ID No. 46; SEQ ID No. 47; SEQ ID No. 93;
SEQ ID No. 101; SEQ ID No. 102; SEQ ID No. 103;
SEQ ID No. 106; SEQ ID No. 107; SEQ ID No. 120;
SEQ ID No. 121; SEQ ID No. 130; SEQ ID No. 135;
SEQ ID No. 140; SEQ ID No. 143; SEQ ID No. 144;
SEQ ID No. 145; SEQ ID No. 158; SEQ ID No. 159;
SEQ ID No. 160; SEQ ID No. 161; SEQ ID No. 192;
SEQ ID No. 193; SEQ ID No. 196; SEQ ID No. 196;
SEQ ID No. 197; SEQ ID No. 198; SEQ ID No. 199;
SEQ ID No. 227; SEQ ID No. 229; SEQ ID No. 236;
SEQ ID No. 236; SEQ ID No. 239; SEQ ID No. 243;
SEQ ID No. 245; SEQ ID No. 264; SEQ ID No. 265;
SEQ ID No. 297; SEQ ID No. 331; SEQ ID No. 333;
SEQ ID No. 359; SEQ ID No. 360; SEQ ID No. 374;
SEQ ID No. 404; SEQ ID No. 405; SEQ ID No. 405;
SEQ ID No. 410; SEQ ID No. 415; SEQ ID No. 415;
SEQ ID No. 416; SEQ ID No. 417; SEQ ID No. 432;
SEQ ID No. 460; SEQ ID No. 461; SEQ ID No. 462;
SEQ ID No. 495; SEQ ID No. 513; SEQ ID No. 515;
SEQ ID No. 566; SEQ ID No. 566; SEQ ID No. 566;

SEQ ID No. 589; SEQ ID No. 613; SEQ ID No. 645;
SEQ ID No. 646; SEQ ID No. 647; SEQ ID No. 652;
SEQ ID No. 653; SEQ ID No. 654; SEQ ID No. 672;
SEQ ID No. 673; SEQ ID No. 674; SEQ ID No. 682;
SEQ ID No. 684; SEQ ID No. 692; SEQ ID No. 700;
SEQ ID No. 725; SEQ ID No. 801; SEQ ID No. 802;
SEQ ID No. 835; SEQ ID No. 836; SEQ ID No. 837;
SEQ ID No. 860; SEQ ID No. 861; SEQ ID No. 862;
SEQ ID No. 863; SEQ ID No. 869; SEQ ID No. 869;
SEQ ID No. 925; SEQ ID No. 964; SEQ ID No. 983 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleotides, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 142; SEQ ID No. 142; SEQ ID No. 169;
SEQ ID No. 256; SEQ ID No. 268; SEQ ID No. 325;
SEQ ID No. 352; SEQ ID No. 366; SEQ ID No. 435;
SEQ ID No. 444; SEQ ID No. 528; SEQ ID No. 529;
SEQ ID No. 530; SEQ ID No. 548; SEQ ID No. 549;
SEQ ID No. 601; SEQ ID No. 602; SEQ ID No. 617;
SEQ ID No. 619; SEQ ID No. 644; SEQ ID No. 745;
SEQ ID No. 971; SEQ ID No. 972; SEQ ID No. 1023 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleic acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 5; SEQ ID No. 12; SEQ ID No. 82;
SEQ ID No. 96; SEQ ID No. 97; SEQ ID No. 98;
SEQ ID No. 99; SEQ ID No. 100; SEQ ID No. 105;
SEQ ID No. 118; SEQ ID No. 136; SEQ ID No. 137;
SEQ ID No. 163; SEQ ID No. 190; SEQ ID No. 204;
SEQ ID No. 259; SEQ ID No. 260; SEQ ID No. 262;
SEQ ID No. 290; SEQ ID No. 300; SEQ ID No. 301;

SEQ ID No. 302; SEQ ID No. 387; SEQ ID No. 427;
SEQ ID No. 434; SEQ ID No. 441; SEQ ID No. 444;
SEQ ID No. 471; SEQ ID No. 595; SEQ ID No. 596;
SEQ ID No. 597; SEQ ID No. 599; SEQ ID No. 600;
SEQ ID No. 605; SEQ ID No. 612; SEQ ID No. 624;
SEQ ID No. 625; SEQ ID No. 650; SEQ ID No. 657;
SEQ ID No. 658; SEQ ID No. 702; SEQ ID No. 703;
SEQ ID No. 704; SEQ ID No. 708; SEQ ID No. 719;
SEQ ID No. 766; SEQ ID No. 767; SEQ ID No. 775;
SEQ ID No. 779; SEQ ID No. 787; SEQ ID No. 788;
SEQ ID No. 794; SEQ ID No. 841; SEQ ID No. 842;
SEQ ID No. 883; SEQ ID No. 884; SEQ ID No. 907;
SEQ ID No. 918; SEQ ID No. 924; SEQ ID No. 928;
SEQ ID No. 929; SEQ ID No. 962; SEQ ID No. 962;
SEQ ID No. 963; SEQ ID No. 969; SEQ ID No. 970;
SEQ ID No. 975; SEQ ID No. 979; SEQ ID No. 995;
SEQ ID No. 1031; SEQ ID No. 1032 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of amino acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 27; SEQ ID No. 41; SEQ ID No. 55;
SEQ ID No. 56; SEQ ID No. 57; SEQ ID No. 59;
SEQ ID No. 62; SEQ ID No. 63; SEQ ID No. 64;
SEQ ID No. 65; SEQ ID No. 119; SEQ ID No. 132;
SEQ ID No. 240; SEQ ID No. 241; SEQ ID No. 277;
SEQ ID No. 278; SEQ ID No. 279; SEQ ID No. 382;
SEQ ID No. 406; SEQ ID No. 428; SEQ ID No. 442;
SEQ ID No. 446; SEQ ID No. 447; SEQ ID No. 453;
SEQ ID No. 454; SEQ ID No. 541; SEQ ID No. 542;
SEQ ID No. 591; SEQ ID No. 608; SEQ ID No. 609;
SEQ ID No. 610; SEQ ID No. 618; SEQ ID No. 648;
SEQ ID No. 649; SEQ ID No. 660; SEQ ID No. 661;
SEQ ID No. 677; SEQ ID No. 717; SEQ ID No. 765;
SEQ ID No. 797; SEQ ID No. 871; SEQ ID No. 875;
SEQ ID No. 920; SEQ ID No. 922; SEQ ID No. 937;

SEQ ID No. 998; SEQ ID No. 1020; SEQ ID No. 1021; SEQ ID No. 1034; SEQ ID No. 1044; SEQ ID No. 1046; SEQ ID No. 1049 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of polypeptides, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 21; SEQ ID No. 21; SEQ ID No. 22; SEQ ID No. 23; SEQ ID No. 24; SEQ ID No. 25; SEQ ID No. 26; SEQ ID No. 75; SEQ ID No. 84; SEQ ID No. 84; SEQ ID No. 86; SEQ ID No. 92; SEQ ID No. 133; SEQ ID No. 151; SEQ ID No. 152; SEQ ID No. 157; SEQ ID No. 179; SEQ ID No. 209; SEQ ID No. 307; SEQ ID No. 326; SEQ ID No. 343; SEQ ID No. 344; SEQ ID No. 345; SEQ ID No. 371; SEQ ID No. 429; SEQ ID No. 519; SEQ ID No. 557; SEQ ID No. 586; SEQ ID No. 587; SEQ ID No. 630; SEQ ID No. 656; SEQ ID No. 706; SEQ ID No. 707; SEQ ID No. 730; SEQ ID No. 751; SEQ ID No. 752; SEQ ID No. 786; SEQ ID No. 847; SEQ ID No. 885; SEQ ID No. 923; SEQ ID No. 978; SEQ ID No. 1039; SEQ ID No. 1048 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of fatty acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 4; SEQ ID No. 15; SEQ ID No. 16; SEQ ID No. 141; SEQ ID No. 173; SEQ ID No. 205; SEQ ID No. 205; SEQ ID No. 206; SEQ ID No. 207; SEQ ID No. 208; SEQ ID No. 312; SEQ ID No. 355; SEQ ID No. 415; SEQ ID No. 550; SEQ ID No. 558; SEQ ID No. 560; SEQ ID No. 561; SEQ ID No. 574; SEQ ID No. 574; SEQ ID No. 577; SEQ ID No. 578; SEQ ID No. 590; SEQ ID No. 614; SEQ ID No. 772; SEQ ID No. 808; SEQ ID No. 809; SEQ ID No. 904;

SEQ ID No. 905; SEQ ID No. 905; SEQ ID No. 933; SEQ ID No. 934; SEQ ID No. 934; SEQ ID No. 936 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the synthesis of the wall, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 87; SEQ ID No. 196; SEQ ID No. 242; SEQ ID No. 269; SEQ ID No. 628; SEQ ID No. 629; SEQ ID No. 634; SEQ ID No. 635; SEQ ID No. 637; SEQ ID No. 638; SEQ ID No. 1019 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the transcription, translation and/or maturation process, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 112; SEQ ID No. 113; SEQ ID No. 332; SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 350; SEQ ID No. 362; SEQ ID No. 363; SEQ ID No. 364; SEQ ID No. 407; SEQ ID No. 451; SEQ ID No. 546; SEQ ID No. 643; SEQ ID No. 744; SEQ ID No. 746; SEQ ID No. 833; SEQ ID No. 868; SEQ ID No. 981; SEQ ID No. 982; SEQ ID No. 1003; SEQ ID No. 1011; SEQ ID No. 1042 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* ribosomal polypeptide or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 114; SEQ ID No. 115; SEQ ID No. 116; SEQ ID No. 328; SEQ ID No. 361; SEQ ID No. 375; SEQ ID No. 445; SEQ ID No. 543; SEQ ID No. 584; SEQ ID No. 585; SEQ ID No. 743; SEQ ID No. 813;

SEQ ID No. 941; SEQ ID No. 942; SEQ ID No. 944;
SEQ ID No. 946; SEQ ID No. 947; SEQ ID No. 948;
SEQ ID No. 950; SEQ ID No. 951; SEQ ID No. 952;
SEQ ID No. 953; SEQ ID No. 954; SEQ ID No. 955;
SEQ ID No. 955; SEQ ID No. 957; SEQ ID No. 958;
SEQ ID No. 960; SEQ ID No. 961; SEQ ID No. 1040;
SEQ ID No. 1041; SEQ ID No. 1043; SEQ ID No. 1063;
SEQ ID No. 1064 and one of their fragments.

Preferably, the invention also relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* transport polypeptide or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 6; SEQ ID No. 50; SEQ ID No. 51;
SEQ ID No. 80; SEQ ID No. 125; SEQ ID No. 126;
SEQ ID No. 128; SEQ ID No. 129; SEQ ID No. 215;
SEQ ID No. 246; SEQ ID No. 248; SEQ ID No. 249;
SEQ ID No. 251; SEQ ID No. 252; SEQ ID No. 253;
SEQ ID No. 255; SEQ ID No. 271; SEQ ID No. 275;
SEQ ID No. 293; SEQ ID No. 309; SEQ ID No. 323;
SEQ ID No. 324; SEQ ID No. 398; SEQ ID No. 401;
SEQ ID No. 449; SEQ ID No. 511; SEQ ID No. 512;
SEQ ID No. 564; SEQ ID No. 565; SEQ ID No. 667;
SEQ ID No. 679; SEQ ID No. 680; SEQ ID No. 711;
SEQ ID No. 712; SEQ ID No. 713; SEQ ID No. 714;
SEQ ID No. 715; SEQ ID No. 730; SEQ ID No. 731;
SEQ ID No. 736; SEQ ID No. 737; SEQ ID No. 738;
SEQ ID No. 870; SEQ ID No. 908; SEQ ID No. 919;
SEQ ID No. 977; SEQ ID No. 987; SEQ ID No. 988;
SEQ ID No. 992; SEQ ID No. 993; SEQ ID No. 994;
SEQ ID No. 1028; SEQ ID No. 1029 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the virulence process, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 20; SEQ ID No. 815; SEQ ID No. 816; SEQ ID No. 898; SEQ ID No. 1059; SEQ ID No. 1060 and one of their fragments.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the secretory system and/or which is secreted, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 758; SEQ ID No. 888; SEQ ID No. 889; SEQ ID No. 890; SEQ ID No. 891; SEQ ID No. 896; SEQ ID No. 897; SEQ ID No. 898 and one of their fragments.

The secreted polypeptides of the present invention, as well as the corresponding nucleotide sequences, may be detected by techniques known to persons skilled in the art, such as for example the techniques using cloning combined with vectors allowing the expression of the said polypeptides fused to export markers such as the *luc* gene for luciferase or the *PhoA* gene for alkaline phosphatase.

Preferably, the invention relates to a polypeptide according to the invention, characterized in that it is a polypeptide specific to *Chlamydiae* or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 22; SEQ ID No. 29; SEQ ID No. 31;
SEQ ID No. 32; SEQ ID No. 34; SEQ ID No. 35;
SEQ ID No. 39; SEQ ID No. 40; SEQ ID No. 43;
SEQ ID No. 48; SEQ ID No. 49; SEQ ID No. 50;
SEQ ID No. 52; SEQ ID No. 53; SEQ ID No. 54;
SEQ ID No. 72; SEQ ID No. 77; SEQ ID No. 78;
SEQ ID No. 87; SEQ ID No. 90; SEQ ID No. 95;
SEQ ID No. 108; SEQ ID No. 110; SEQ ID No. 111;
SEQ ID No. 122; SEQ ID No. 123; SEQ ID No. 124;
SEQ ID No. 127; SEQ ID No. 138; SEQ ID No. 144;
SEQ ID No. 146; SEQ ID No. 153; SEQ ID No. 155;
SEQ ID No. 164; SEQ ID No. 166; SEQ ID No. 175;

SEQ ID No. 182; SEQ ID No. 184; SEQ ID No. 186;
SEQ ID No. 187; SEQ ID No. 188; SEQ ID No. 202;
SEQ ID No. 210; SEQ ID No. 247; SEQ ID No. 258;
SEQ ID No. 266; SEQ ID No. 267; SEQ ID No. 270;
SEQ ID No. 273; SEQ ID No. 274; SEQ ID No. 295;
SEQ ID No. 296; SEQ ID No. 305; SEQ ID No. 306;
SEQ ID No. 309; SEQ ID No. 318; SEQ ID No. 319;
SEQ ID No. 322; SEQ ID No. 326; SEQ ID No. 342;
SEQ ID No. 357; SEQ ID No. 376; SEQ ID No. 379;
SEQ ID No. 380; SEQ ID No. 388; SEQ ID No. 390;
SEQ ID No. 400; SEQ ID No. 431; SEQ ID No. 433;
SEQ ID No. 438; SEQ ID No. 443; SEQ ID No. 456;
SEQ ID No. 457; SEQ ID No. 458; SEQ ID No. 464;
SEQ ID No. 468; SEQ ID No. 470; SEQ ID No. 473;
SEQ ID No. 486; SEQ ID No. 489; SEQ ID No. 497;
SEQ ID No. 501; SEQ ID No. 503; SEQ ID No. 504;
SEQ ID No. 508; SEQ ID No. 512; SEQ ID No. 521;
SEQ ID No. 522; SEQ ID No. 523; SEQ ID No. 524;
SEQ ID No. 533; SEQ ID No. 535; SEQ ID No. 536;
SEQ ID No. 537; SEQ ID No. 538; SEQ ID No. 539;
SEQ ID No. 540; SEQ ID No. 554; SEQ ID No. 563;
SEQ ID No. 572; SEQ ID No. 579; SEQ ID No. 595;
SEQ ID No. 603; SEQ ID No. 604; SEQ ID No. 606;
SEQ ID No. 607; SEQ ID No. 615; SEQ ID No. 616;
SEQ ID No. 622; SEQ ID No. 641; SEQ ID No. 642;
SEQ ID No. 659; SEQ ID No. 668; SEQ ID No. 670;
SEQ ID No. 693; SEQ ID No. 695; SEQ ID No. 696;
SEQ ID No. 699; SEQ ID No. 703; SEQ ID No. 704;
SEQ ID No. 716; SEQ ID No. 726; SEQ ID No. 728;
SEQ ID No. 739; SEQ ID No. 742; SEQ ID No. 747;
SEQ ID No. 750; SEQ ID No. 751; SEQ ID No. 755;
SEQ ID No. 757; SEQ ID No. 759; SEQ ID No. 761;
SEQ ID No. 762; SEQ ID No. 763; SEQ ID No. 764;
SEQ ID No. 773; SEQ ID No. 780; SEQ ID No. 781;
SEQ ID No. 789; SEQ ID No. 800; SEQ ID No. 803;
SEQ ID No. 804; SEQ ID No. 818; SEQ ID No. 820;
SEQ ID No. 822; SEQ ID No. 823; SEQ ID No. 824;
SEQ ID No. 827; SEQ ID No. 828; SEQ ID No. 839;
SEQ ID No. 849; SEQ ID No. 850; SEQ ID No. 851;

SEQ ID No. 852; SEQ ID No. 855; SEQ ID No. 856;
SEQ ID No. 857; SEQ ID No. 858; SEQ ID No. 859;
SEQ ID No. 860; SEQ ID No. 861; SEQ ID No. 862;
SEQ ID No. 863; SEQ ID No. 865; SEQ ID No. 868;
SEQ ID No. 869; SEQ ID No. 870; SEQ ID No. 871;
SEQ ID No. 872; SEQ ID No. 873; SEQ ID No. 874;
SEQ ID No. 875; SEQ ID No. 877; SEQ ID No. 878;
SEQ ID No. 880; SEQ ID No. 882; SEQ ID No. 884;
SEQ ID No. 886; SEQ ID No. 893; SEQ ID No. 901;
SEQ ID No. 906; SEQ ID No. 910; SEQ ID No. 912;
SEQ ID No. 915; SEQ ID No. 916; SEQ ID No. 917;
SEQ ID No. 926; SEQ ID No. 929; SEQ ID No. 933;
SEQ ID No. 965; SEQ ID No. 967; SEQ ID No. 968;
SEQ ID No. 984; SEQ ID No. 986; SEQ ID No. 989;
SEQ ID No. 990; SEQ ID No. 996; SEQ ID No. 997;
SEQ ID No. 1001; SEQ ID No. 1002; SEQ ID No. 1013;
SEQ ID No. 1016; SEQ ID No. 1031; SEQ ID No. 1033;
SEQ ID No. 1035; SEQ ID No. 1049; SEQ ID No. 1051;
SEQ ID No. 1052; SEQ ID No. 1054; SEQ ID No. 1056;
SEQ ID No. 1057; SEQ ID No. 1058; SEQ ID No. 1062;
SEQ ID No. 1070; SEQ ID No. 1071; SEQ ID No. 1073 and one of their fragments.

In general, in the present invention, the functional group to which a polypeptide of the invention belongs, as well as its corresponding nucleotide sequence, may be determined either by comparative analogy with sequences already known, or by the use of standard techniques of biochemistry, of cytology combined with the techniques of genetic engineering such as immunoaffinity, localization by immunolabelling, differential extraction, measurement of enzymatic activity, study of the activity inducing or repressing expression or the study of expression in E. coli.

It is clearly understood, on the one hand, that, in the present invention, the nucleotide sequences (ORF) and the amino acid sequences (SEQ ID No. 2 to SEQ ID No. 1076) which are listed by functional group, are not exhaustive within the group

considered. Moreover, it is also clearly understood that, in the present invention, a nucleotide sequence (ORF) or an amino acid sequence mentioned within a given functional group may also be part of another group taking into account, for example, the interrelationship between the groups listed. Accordingly, and as an example of this interrelationship, an exported and/or secreted polypeptide as well as its coding nucleotide sequence may also be involved in the *Chlamydia trachomatis* virulence process by modifying the defence mechanism of the infected host cell, or a transmembrane polypeptide or its coding nucleotide sequence is also part of the polypeptides or coding nucleotide sequences of the cellular envelope.

The subject of the present invention is also the nucleotide and/or polypeptide sequences according to the invention, characterized in that the said sequences are recorded on a medium, called recording medium, whose type and nature facilitate the reading, the analysis and the exploitation of the said sequences. These media may of course also contain other information extracted from the present invention, such as in particular the analogies with already known sequences, such as those mentioned in Table 1 of the present description, and/or may contain, in addition, information relating to the nucleotide and/or polypeptide sequences of other microorganisms so as to facilitate the comparative analysis and the exploitation of the results obtained.

Among these recording media, computer-readable media, such as magnetic, optical, electrical and hybrid media such as, for example, floppy disks, CD-ROMs or recording cassettes, are preferred in particular.

The invention also relates to nucleotide sequences which can be used as primer or probe, characterized in that the said sequences are chosen from the nucleotide sequences according to the invention.

The invention relates, in addition, to the use of a nucleotide sequence according to the invention, as primer or probe, for the detection and/or amplification of nucleic acid sequences.

The nucleotide sequences according to the invention may thus be used to amplify nucleotide sequences, in particular by the PCR technique (polymerase chain reaction) (Erlich, 1989; Innis et al., 1990; Rolfs et al., 1991, and White et al., 1997).

These oligodeoxyribonucleotide or oligoribonucleotide primers are advantageously at least 8 nucleotides, preferably at least 12 nucleotides, and still more preferably at least 20 nucleotides long.

Other techniques for amplifying the target nucleic acid may be advantageously used as alternatives to PCR.

The nucleotide sequences of the invention, in particular the primers according to the invention, may also be used in other methods for amplifying a target nucleic acid, such as:

- the TAS (Transcription-based Amplification System) technique described by Kwoh et al. in 1989;
- the 3SR (Self-Sustained Sequence Replication) technique described by Guatelli et al. in 1990;
- the NASBA (Nucleic Acid Sequence Based Amplification) technique described by Kievitis et al. in 1991;
- the SDA (Strand Displacement Amplification) technique (Walker et al., 1992);
- the TMA (Transcription Mediated Amplification) technique.

The polynucleotides of the invention may also be used in techniques for amplifying or for modifying the nucleic acid serving as probe, such as:

- the LCR (Ligase Chain Reaction) technique described by Landegren et al. in 1988 and perfected by Barany et al. in 1991, which uses a thermostable ligase;

- the RCR (Repair Chain Reaction) technique described by Segev in 1992;
- the CPR (Cycling Probe Reaction) technique described by Duck et al. in 1990;
- the Q-beta-replicase amplification technique described by Miele et al. in 1983 and perfected in particular by Chu et al. in 1986, Lizardi et al. in 1988, and then by Burg et al. as well as by Stone et al. in 1996.

The invention also relates to the nucleotide sequences of fragments which can be obtained by amplification with the aid of at least one primer according to the invention.

In the case where the target polynucleotide to be detected is possibly an RNA, for example an mRNA, it will be possible to use, prior to the use of an amplification reaction with the aid of at least one primer according to the invention or to the use of a method of detection with the aid of at least one probe of the invention, a reverse transcriptase-type enzyme so as to obtain a cDNA from the RNA contained in the biological sample. The cDNA obtained will then serve as target for the primer(s) or the probe(s) used in the amplification or detection method according to the invention.

The detection probe will be chosen so that it hybridizes with the target sequence or the amplicon generated from the target sequence. Such a detection probe will advantageously have as sequence a sequence of at least 12 nucleotides, in particular of at least 20 nucleotides, and preferably at least 100 nucleotides.

The invention also comprises the nucleotide sequences which can be used as probe or primer according to the invention, characterized in that they are labelled with a radioactive compound or with a nonradioactive compound.

The nonlabelled nucleotide sequences may be used directly as probes or primers; however, the

sequences are generally labelled with a radioactive element (^{32}P, ^{35}S, ^{3}H, ^{125}I) or with a nonradioactive molecule (biotin, acetylaminofluorene, digoxigenin, 5-bromo-deoxyuridine, fluorescein) so as to obtain probes which can be used in numerous applications.

Examples of nonradioactive labelling of nucleotide sequences are described, for example, in French patent No. 78,10975 or by Urdea et al. or by Sanchez-Pescador et al. in 1988.

In the latter case, one of the labelling methods described in patents FR-2 422 956 and FR-2 518 755 may also be used.

The invention also relates to the nucleotide sequences of fragments which can be obtained by hybridization with the aid of at least one probe according to the invention.

The hybridization technique may be performed in various ways (Matthews et al., 1988). The most common method consists in immobilizing the nucleic acid extracted from *C. trachomatis* cells on a support (such as nitrocellulose, nylon, polystyrene) and in incubating, under well-defined conditions, the target nucleic acid immobilized with the probe. After hybridization, the excess probe is removed and the hybrid molecules formed are detected by the appropriate method (measurement of the radioactivity, of the fluorescence or of the enzymatic activity linked to the probe).

The invention also comprises the nucleotide sequences according to the invention, characterized in that they are covalently or noncovalently immobilized on a support.

According to another advantageous embodiment of the nucleic sequences according to the invention, the latter may be used immobilized on a support and may thus serve to capture, through specific hybridization, the target nucleic acid obtained from the biological sample to be tested. If necessary, the solid support is separated from the sample and the hybridization complex

formed between the so-called capture probe and the target nucleic acid is then detected by means of a second probe, called detection probe, labelled with an easily detectable element.

The nucleotide sequences according to the invention may also be used in new analytical systems, DNA chips, which allow sequencing, the study of mutations and of the expression of genes, and which are currently of interest given their very small size and their high capacity in terms of number of analyses.

The principle of the operation of these chips is based on molecular probes, most often oligonucleotides, which are attached onto a miniaturized surface, generally of the order of a few square centimetres. During an analysis, a sample containing fragments of a target nucleic acid to be analysed, for example DNA or RNA labelled, for example, after amplification, is deposited onto the DNA chip in which the support has been coated beforehand with probes. Bringing the labelled target sequences into contact with the probes leads to the formation, through hybridization, of a duplex according to the rule of pairing defined by J.D. Watson and F. Crick. After a washing step, analysis of the surface of the chip allows the effective hybridizations to be located by means of the signals emitted by the labels tagging the target. A hybridization fingerprint results from this analysis which, by appropriate computer processing, will make it possible to determine information such as the presence of specific fragments in the sample, the determination of sequences and the presence of mutations.

The chip consists of a multitude of molecular probes, precisely organized on a solid support whose surface is miniaturized. It is at the centre of a system where other elements (imaging system, microcomputer) allow the acquisition and interpretation of a hybridization fingerprint.

The hybridization supports are provided in the

form of flat or porous surfaces (pierced with wells) composed of various materials. The choice of a support is determined by its physicochemical properties, or more precisely, by the relationship between the latter and the conditions under which the support will be placed during the synthesis or the attachment of the probes or during the use of the chip. It is therefore necessary, before considering the use of a particular support (R.S. Matson et al., 1994), to consider characteristics such as its stability to pH, its physical strength, its reactivity and its chemical stability as well as its capacity to nonspecifically bind nucleic acids. Materials such as glass, silicon and polymers are commonly used. Their surface is, in a first step, called "functionalization", made reactive towards the groups which it is desired to attach thereon. After the functionalization, so-called spacer molecules are grafted onto the activated surface. Used as intermediates between the surface and the probe, these molecules of variable size render unimportant the surface properties of the supports, which often prove to be problematic for the synthesis or the attachment of the probes and for the hybridization.

Among the hybridization supports, there may be mentioned glass which is used, for example, in the method of in situ synthesis of oligonucleotides by photochemical addressing developed by the company Affymetrix (E.L. Sheldon, 1993), the glass surface being activated by silane. Genosensor Consortium (P. Mérel, 1994) also uses glass slides carrying wells 3 mm apart, this support being activated with epoxysilane.

Polymers or silicon may also be mentioned among these hybridization supports. For example, the Andrein Mirzabekov team has developed a chip consisting of polyacrylamide squares polymerized on a silanized glass surface (G. Yershov et al., 1996). Several teams use silicon, in particular the IFOS laboratory of Ecole Centrale of Lyon which uses a silicon semiconductor

substrate which is p-doped by introducing it into its crystalline structure atoms whose valency is different from that of silicon. Various types of metals, in particular gold and platinum, may also be used as support (Genosensor Consortium (K. Beattie et al., 1993)).

The probes according to the invention may be synthesized directly in situ on the supports of the DNA chips. This in situ synthesis may be carried out by photochemical addressing (developed by the company Affymax (Amsterdam, Holland) and exploited industrially by its subsidiary Affymetrix (United States)) or based on the VLSIPS (very large scale immobilized polymer synthesis) technology (S.P.A. Fodor et al., 1991) which is based on a method of photochemically directed combinatory synthesis and the principle of which combines solid-phase chemistry, the use of photolabile protecting groups and photolithography.

The probes according to the invention may be attached to the DNA chips in various ways such as electrochemical addressing, automated addressing or the use of probe printers (T. Livache et al., 1994; G. Yershov et al., 1996; J. Derisi et al., 1996, and S. Borman, 1996).

The revealing of the hybridization between the probes of the invention, deposited or synthesized in situ on the supports of the DNA chips, and the sample to be analysed, may be determined, for example, by measurement of fluorescent signals, by radioactive counting or by electronic detection.

The use of fluorescent molecules such as fluorescein constitutes the most common method of labelling the samples. It allows direct or indirect revealing of the hybridization and allows the use of various fluorochromes.

Affymetrix currently provides an apparatus or a scanner designed to read its Gene Chip™ chips. It makes it possible to detect the hybridizations by scanning the surface of the chip in confocal microscopy

(R.J. Lipshutz et al., 1995). Other methods of detecting fluorescent signals have been tested: coupling of an epifluorescence microscope and a CCD camera (G. Yershov et al., 1996), the use of an optical fibre collecting system (E.L. Sheldon, 1993). A conventional method consists in carrying out an end labelling, with phosphorus 32, of the target sequences, by means of an appropriate apparatus, the Phosphorimager (marketed by Molecular Dynamics). The electronic detection is based on the principle that the hybridization of two nucleic acid molecules is accompanied by physical phenomena which can be quantified under certain conditions (system developed by Ecole Centrale of Lyon and called GEN-FET (GEN field effect transistor)). Genosensor Consortium and the company Beckman Instruments who are developing an electronic chip or Permittivity Chips™ may also be mentioned (K. Beattie et al., 1993).

The nucleotide sequences according to the invention may thus be used in DNA chips to carry out the analysis of mutations. This analysis is based on the production of chips capable of analysing each base of a nucleotide sequence according to the invention. It is possible, in particular to this end, to use the microsequencing techniques on a DNA chip. The mutations are detected by extending immobilized primers which hybridize to the template of sequences analysed, just at the position adjacent to that of the mutated nucleotide to be detected. A single-stranded template, RNA or DNA, of the sequences to be analysed will be advantageously prepared according to conventional methods, from products amplified according to PCR-type techniques. The templates of single-stranded DNA, or of RNA thus obtained are then deposited on the DNA chip, under conditions allowing their specific hybridization to the immobilized primers. A thermostable polymerase, for example Tth or T7 DNA polymerase, specifically extends the 3′ end of the immobilized primer with a labelled nucleotide analogue complementary to the

nucleotide at the position of the variable site. For example a thermal cycling is performed in the presence of fluorescent dideoxyribonucleotides. The experimental conditions will be adapted in particular to the chips used, to the immobilized primers, to the polymerases used and to the labelling system chosen. One advantage of microsequencing, compared with techniques based on the hybridization of probes, is that it makes it possible to identify all the variable nucleotides with optimal discrimination under homogeneous reaction conditions; used on DNA chips, it allows optimal resolution and specificity for the routine and industrial detection of mutations in multiplex.

The nucleotide sequences according to the invention may also be used in DNA chips to carry out the analysis of the expression of the *Chlamydia trachomatis* genes. This analysis of the expression of *Chlamydia trachomatis* genes is based on the use of chips where probes of the invention, chosen for their specificity to characterize a given gene, are present (D.J. Lockhart et al., 1996; D.D. Shoemaker et al., 1996). For the methods of analysis of gene expression using the DNA chips, reference may, for example, be made to the methods described by D.J. Lockhart et al. (1996) and Sosnowsky et al. (1997) for the synthesis of probes in situ or for the addressing and the attachment of previously synthesized probes. The target sequences to be analysed are labelled and in general fragmented into sequences of about 50 to 100 nucleotides before being hybridized onto the chip. After washing as described, for example, by D.J. Lockhart et al. (1996) and application of different electric fields (Sosnowsky et al., 1997), the labelled compounds are detected and quantified, the hybridizations being carried out at least in duplicate. Comparative analyses of the signal intensities obtained with respect to the same probe for different samples and/or for different probes with the same sample, determine the differential expression of RNA or of DNA derived from the sample.

The nucleotide sequences according to the invention may, in addition, be used in DNA chips where other nucleotide probes specific for other microorganisms are also present, and may allow the carrying out of a serial test allowing rapid identification of the presence of a microorganism in a sample.

Accordingly, the subject of the invention is also the nucleotide sequences according to the invention, characterized in that they are immobilized on a support of a DNA chip.

The DNA chips, characterized in that they contain at least one nucleotide sequence according to the invention, immobilized on the support of the said chip, also form part of the invention.

The said chips will preferably contain several probes or nucleotide sequences of the invention of different length and/or corresponding to different genes so as to identify, with greater certainty, the specificity of the target sequences or the desired mutation in the sample to be analysed.

Accordingly, the analyses carried out by means of primers and/or probes according to the invention, immobilized on supports such as DNA chips, will make it possible, for example, to identify, in samples, mutations linked to variations such as intraspecies variations. These variations may be correlated or associated with pathologies specific to the variant identified and will make it possible to select the appropriate treatment.

The invention thus comprises a DNA chip according to the invention, characterized in that it contains, in addition, at least one nucleotide sequence of a microorganism different from *Chlamydia trachomatis,* immobilized on the support of the said chip; preferably, the different microorganism will be chosen from an associated microorganism, a bacterium of the *Chlamydia* family, and a variant of the species *Chlamydia trachomatis*.

Another subject of the present invention is a vector for the cloning and/or the expression of a sequence, characterized in that it contains a nucleotide sequence according to the invention.

Among the said vectors according to the invention, the vectors containing a nucleotide sequence encoding a polypeptide of the cellular, preferably outer, envelope of *Chlamydia trachomatis* or one of its fragments, are preferred.

Among the said vectors according to the invention, the vectors containing a nucleotide sequence encoding a *Chlamydia trachomatis* secreted polypeptide or one of its fragments or encoding a transport polypeptide, a ribosomal polypeptide or a polypeptide involved in secretion, transcription, translation, maturation of proteins, a polypeptide involved in the synthesis of the wall, a polypeptide involved in the virulence, a polypeptide involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, a polypeptide involved in the metabolism of nucleotides, of amino acids, of nucleic acids or of fatty acids of *Chlamydia trachomatis* or one of their fragments, or a polypeptide specific to *Chlamydiae*, are also preferred.

The vectors according to the invention, characterized in that they comprise the elements allowing the expression and/or the secretion of the said nucleotide sequences in a given host cell, also form part of the invention.

The vector should, in this case, comprise a promoter, signals for initiation and for termination of translation, as well as appropriate regions for regulation of transcription. It should be capable of being stably maintained in the host cell and may optionally possess particular signals specifying the secretion of the translated protein. These different elements are chosen according to the host cell used. To this effect, the nucleotide sequences according to the invention may be inserted into autonomously-replicating

vectors within the chosen host, or integrative vectors in the chosen host.

Such vectors will be prepared according to the methods commonly used by persons skilled in the art, and the clones resulting therefrom may be introduced into an appropriate host by standard methods, such as for example lipofection, electroporation and heat shock.

The vectors according to the invention are, for example, vectors of plasmid or viral origin.

A preferred vector for the expression of the polypeptides of the invention consists of the pET-type plasmid vectors (Promega).

These vectors are useful for transforming host cells so as to clone or express the nucleotide sequences of the invention.

The invention also comprises the host cells transformed by a vector according to the invention.

These cells may be obtained by introducing into host cells a nucleotide sequence inserted into a vector as defined above, and then culturing the said cells under conditions allowing the replication and/or the expression of the transfected nucleotide sequence.

The host cell may be chosen from eukaryotic or prokaryotic systems, such as for example bacterial cells (Olins and Lee, 1993), but also yeast cells (Buckholz, 1993), as well as animal cells, in particular cultures of mammalian cells (Edwards and Aruffo, 1993), and in particular Chinese hamster ovary (CHO) cells, but also insect cells in which methods using baculoviruses for example may be used (Luckow, 1993).

A preferred host cell for the expression of the proteins of the invention consists of prokaryotic cells, such as Gram$^-$ bacteria.

A further preferred host cell according to the invention is a bacterium belonging to the *Chlamydia* family, more preferably belonging to the species *Chlamydia trachomatis* or chosen from a microorganism

associated with the species *Chlamydia trachomatis*.

The invention also relates to the animals, except humans, comprising one of the said transformed cells according to the invention.

The production of transgenic animals according to the invention overexpressing one or more of the *Chlamydia trachomatis* genes will be preferably carried out on rats, mice or rabbits according to methods well known to persons skilled in the art such as viral or nonviral transfections. The transgenic animals over-expressing one or more of the said genes may be obtained by transfection of multiple copies of the said genes under the control of a powerful promoter of a ubiquitous nature, or which is selective for one type of tissue. The transgenic animals may also be obtained by homologous recombination on embryonic stem cells, transfer of these stem cells to embryos, selection of the chimeras affected at the level of the reproductive lines, and growth of the said chimeras.

The transformed cells as well as the transgenic animals according to the invention can be used in methods of preparing the recombinant polypeptide.

It is now possible to produce recombinant polypeptides in a relatively large quantity by genetic engineering using the cells transformed with expression vectors according to the invention or using transgenic animals according to the invention.

The methods of preparing a polypeptide of the invention in recombinant form, characterized in that they use a vector and/or a cell transformed with a vector according to the invention and/or a transgenic animal comprising one of the said transformed cells according to the invention, are themselves included in the present invention.

Among the said methods of preparing a polypeptide of the invention in recombinant form, the methods of preparation using a vector, and/or a cell transformed with the said vector and/or a transgenic animal comprising one of the said transformed cells,

containing a nucleotide sequence encoding a polypeptide of the cellular envelope of *Chlamydia trachomatis* or one of its fragments, more preferably encoding a polypeptide of the outer cellular envelope of *Chlamydia trachomatis* or one of its fragment, are preferred.

Among the said methods of preparing a polypeptide of the invention in recombinant form, the methods of preparation using a vector, and/or a cell transformed with the said vector and/or a transgenic animal comprising one of the said transformed cells, containing a nucleotide sequence encoding a *Chlamydia trachomatis* secreted polypeptide or one of its fragments or encoding a transport polypeptide, a ribosomal polypeptide or a polypeptide involved in secretion, transcription, translation, maturation of proteins, a polypeptide involved in the synthesis of the wall, a polypeptide involved in the virulence, a polypeptide involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, a polypeptide involved in the metabolism of nucleotides, of amino acids, of nucleic acids or of fatty acids of *Chlamydia trachomatis* or one of their fragments, or a polypeptide specific to *Chlamydiae*, are also preferred.

The recombinant polypeptides obtained as indicated above may be provided either in glycosylated or nonglycosylated form and may or may not have the natural tertiary structure.

A preferred variant consists in producing a recombinant polypeptide fused to a "carrier" protein (chimeric protein). The advantage of this system is that it allows a stabilization and a reduction in proteolysis of the recombinant product, an increase in solubility during renaturation in vitro and/or a simplification of purification when the fusion partner has affinity for a specific ligand.

More particularly, the invention relates to a method of preparing a polypeptide of the invention comprising the following steps:

a) culture of the transformed cells under conditions allowing the expression of a recombinant polypeptide having a nucleic acid sequence according to the invention;

b) where appropriate, recovery of the said recombinant polypeptide.

When the method of preparing a polypeptide of the invention uses a transgenic animal according to the invention, the recombinant polypeptide is then extracted from the said animal.

The subject of the invention is also a polypeptide capable of being obtained by a method of the invention as described above.

The invention also comprises a method of preparing a synthetic polypeptide, characterized in that it uses an amino acid sequence of polypeptides according to the invention.

The invention also relates to a synthetic polypeptide obtained by a method according to the invention.

Polypeptides according to the invention may also be prepared by conventional techniques in the field of peptide synthesis. This synthesis may be carried out in a homogeneous solution or on a solid phase.

For example, the synthesis technique in a homogeneous solution described by Houbenweyl in 1974 may be used.

This method of synthesis consists in successively condensing, in pairs, the successive amino acids in the required order, or in condensing amino acids and fragments previously formed and already containing several amino acids in the appropriate order, or alternatively several fragments thus previously prepared, it being understood that care will have been taken to protect beforehand all the reactive functional groups carried by these amino acids or fragments, with the exception of the amine functional groups of one and the carboxyl functional groups of the

other or vice versa, which should normally take part in the formation of the peptide bonds, in particular after activation of the carboxyl functional group, according to methods well known in peptide synthesis.

According to another preferred technique of the invention, the one described by Merrifield is used.

To manufacture a peptide chain according to the Merrifield method, a highly porous polymer resin is used, onto which the first C-terminal amino acid of the chain is attached. This amino acid is attached onto a resin via its carboxyl group and its amine functional group is protected. The amino acids which will constitute the peptide chain are thus attached, one after another, onto the amine group, each time deprotected beforehand, of the portion of the peptide chain already formed, and which is attached to the resin. When the entire peptide chain desired is formed, the protecting groups are removed from the various amino acids constituting the peptide chain and the peptide is detached from the resin with the aid of an acid.

The invention relates, in addition, to hybrid polypeptides having at least one polypeptide or one of its fragments according to the invention, and a sequence of a polypeptide capable of eliciting an immune response in humans or animals.

Advantageously, the antigenic determinant is such that it is capable of eliciting a humoral and/or cellular response.

Such a determinant may comprise a polypeptide or one of its fragments according to the invention, in glycosylated form, used in order to obtain immunogenic compositions capable of inducing the synthesis of antibodies directed against multiple epitopes. The said polypeptides or their glycosylated fragments also form part of the invention.

These hybrid molecules may consist, in part, of a carrier molecule for polypeptides or for their fragments according to the invention, combined with a

portion which may be immunogenic, in particular an epitope of the diphtheria toxin, the tetanus toxin, a hepatitis B virus surface antigen (patent FR 79 21811), the poliomyelitis virus VP1 antigen or any other viral or bacterial toxin or antigen.

The methods of synthesizing the hybrid molecules include the methods used in genetic engineering to construct hybrid nucleotide sequences encoding the desired polypeptide sequences. Reference may be advantageously made, for example, to the technique for producing genes encoding fusion proteins described by Minton in 1984.

The said hybrid nucleotide sequences encoding a hybrid polypeptide as well as the hybrid polypeptides according to the invention, characterized in that they are recombinant polypeptides obtained by the expression of the said hybrid nucleotide sequences, also form part of the invention.

The invention also comprises the vectors characterized in that they contain one of the said hybrid nucleotide sequences. The host cells transformed by the said vectors, the transgenic animals comprising one of the said transformed cells as well as the methods of preparing recombinant polypeptides using the said vectors, the said transformed cells and/or the said transgenic animals of course also form part of the invention.

The polypeptides according to the invention, the antibodies according to the invention described below and the nucleotide sequences according to the invention may advantageously be used in *in vitro* and/or *in vivo* methods for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis*, in a biological sample (biological tissue or fluid) which is likely to contain them. These methods, depending on the specificity of the polypeptides, of the antibodies and of the nucleotide sequences according to the invention which will be used, may in particular detect and/or identify

the bacterial variants belonging to the species *Chlamydia trachomatis* as well as the associated microorganisms capable of being detected by the polypeptides, the antibodies and the nucleotide sequences according to the invention which will be chosen. It may, for example, be advantageous to choose a polypeptide, an antibody or a nucleotide sequence according to the invention, which is capable of detecting any bacterium of the *Chlamydia* family by choosing a polypeptide, an antibody and/or a nucleotide sequence according to the invention which is specific to the family or, on the contrary, it will be most particularly advantageous to target a variant of the species *Chlamydia trachomatis*, which is responsible, for example, for the induction or the worsening of pathologies specific to the targeted variant, by choosing a polypeptide, an antibody and/or a nucleotide sequence according to the invention which is specific to the said variant.

The polypeptides according to the invention may advantageously be used in a method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, in a biological sample (biological tissue or fluid) which is likely to contain them, characterized in that it comprises the following steps:

a) bringing this biological sample into contact with a polypeptide or one of its fragments according to the invention (under conditions allowing an immunological reaction between the said polypeptide and the antibodies which may be present in the biological sample);

b) detecting the antigen-antibody complexes which may be formed.

Preferably, the biological sample consists of a fluid, for example a human or animal serum, blood or biopsies.

Any conventional procedure may be used to carry out such a detection of the antigen-antibody complexes

which may be formed.

By way of example, a preferred method uses immunoenzymatic procedures based on the ELISA technique, immunofluorescence procedures or radio-immunological procedures (RIA), and the like.

Accordingly, the invention also relates to the polypeptides according to the invention, labelled with the aid of a suitable label such as a label of the enzymatic, fluorescent or radioactive type.

Such methods comprise, for example, the following steps:

- deposition of defined quantities of a polypeptide composition according to the invention into the wells of a microtitre plate,
- introduction, into the said wells, of increasing dilutions of serum, or of a different biological sample as defined above, which has to be analysed,
- incubation of the microplate,
- introduction, into the wells of the microtitre plate, of labelled antibodies directed against human or animal immunoglobulins, these antibodies having been labelled with the aid of an enzyme selected from those which are capable of hydrolyzing a substrate, thereby modifying the absorption of the radiation of the latter, at least at a defined wavelength, for example at 550 nm,
- detection, by comparison with a control, of the quantity of substrate hydrolyzed.

The invention also relates to a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

- a polypeptide according to the invention,
- where appropriate, the reagents for constituting the medium appropriate for the immunological or specific reaction,
- the reagents allowing the detection of the antigen-antibody complexes produced by the immuno-

logical reaction between the polypeptide(s) of the invention and the antibodies which may be present in the biological sample, it being possible for these reagents also to carry a label, or to be capable of being recognized in turn by a labelled reagent, more particularly in the case where the polypeptide according to the invention is not labelled,

- where appropriate, a reference biological sample (negative control) free of antibodies recognized by a polypeptide according to the invention,
- where appropriate, a reference biological sample (positive control) containing a predetermined quantity of antibodies recognized by a polypeptide according to the invention.

The polypeptides according to the invention make it possible to prepare monoclonal or polyclonal antibodies characterized in that they specifically recognize the polypeptides according to the invention. The monoclonal antibodies may advantageously be prepared from hybridomas according to the technique described by Kohler and Milstein in 1975. The polyclonal antibodies may be prepared, for example, by immunizing an animal, in particular a mouse, with a polypeptide according to the invention combined with an immune response adjuvant, and then purifying the specific antibodies contained in the serum of the immunized animals on an affinity column onto which the polypeptide which served as antigen has previously been attached. The polyclonal antibodies according to the invention may also be prepared by purifying, on an affinity column onto which a polypeptide according to the invention has previously been attached, the antibodies contained in the serum of patients infected with a bacterium belonging to the species *Chlamydia trachomatis*.

The subject of the invention is also monoclonal or polyclonal antibodies or their fragments, or chimeric antibodies, characterized in that they are capable of specifically recognizing a polypeptide or

one of its fragments according to the invention.

The antibodies of the invention may also be labelled in the same manner as described above for the nucleic probes of the invention such as an enzymatic, fluorescent or radioactive type labelling.

The invention relates, in addition, to a method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it comprises the following steps:

a) bringing the biological sample (biological tissue or fluid) into contact with a mono- or polyclonal antibody according to the invention (under conditions allowing an immunological reaction between the said antibodies and the polypeptides of the bacterium belonging to the species *Chlamydia trachomatis* or to an associated microorganism which may be present in the biological sample);

b) detecting the antigen-antibody complex which may be formed.

Also falling within the scope of the invention is a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

- a polyclonal or monoclonal antibody according to the invention, labelled where appropriate;
- where appropriate, a reagent for constituting the medium appropriate for carrying out the immunological reaction;
- a reagent allowing the detection of the antigen-antibody complexes produced by the immunological reaction, it being possible for this reagent also to carry a label, or to be capable of being recognized in turn by a labelled reagent, more particularly in the case where the said monoclonal or polyclonal antibody is not labelled;
- where appropriate, reagents for carrying out the

lysis of the cells in the sample tested.

The principle of the DNA chip which was explained above may also be used to produce protein "chips" on which the support has been coated with a polypeptide or an antibody according to the invention in place of the DNA. These protein "chips" make it possible, for example, to analyse the biomolecular interactions (BIA) induced by the affinity capture of target analytes onto a support coated, for example, with proteins, by surface plasma resonance (SPR). Reference may be made, for example, to the techniques for coupling proteins onto a solid support which are described in EP 524 800 or to the methods describing the use of biosensor-type protein chips such as the BIAcore-type technique (Pharmacia) (Arlinghaus et al., 1997, Krone et al., 1997, Chatelier et al., 1995). These polypeptides or antibodies according to the invention, capable of specifically binding antibodies or polypeptides derived from the sample to be analysed, may thus be used in protein chips for the detection and/or the identification of proteins in samples. The said protein chips may in particular be used for infectious diagnosis and may preferably contain, per chip, several polypeptides and/or antibodies of the invention of different specificity, and/or polypeptides and/or antibodies capable of recognizing microorganisms different from *Chlamydia trachomatis*.

Accordingly, the subject of the present invention is also the polypeptides and the antibodies according to the invention, characterized in that they are immobilized on a support, in particular of a protein chip.

The protein chips, characterized in that they contain at least one polypeptide or one antibody according to the invention immobilized on the support of the said chip, also form part of the invention.

The invention comprises, in addition, a protein chip according to the invention, characterized in that it contains, in addition, at least one polypeptide of a

microorganism different from *Chlamydia trachomatis* or at least one antibody directed against a compound of a microorganism different from *Chlamydia trachomatis*, immobilized on the support of the said chip.

The invention also relates to a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, or for the detection and/or the identification of a microorganism characterized in that it comprises a protein chip according to the invention.

The subject of the present invention is also a method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it uses a nucleotide sequence according to the invention.

More particularly, the invention relates to a method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it comprises the following steps:

a) where appropriate, isolation of the DNA from the biological sample to be analysed, or optionally production of a cDNA from the RNA in the biological sample;

b) specific amplification of the DNA of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism with the aid of at least one primer according to the invention;

c) detection of the amplification products.

These may be detected, for example, by the molecular hybridization technique using a nucleic probe according to the invention. This probe will be advantageously labelled with a nonradioactive (cold probe) or radioactive element.

For the purposes of the present invention, "DNA in the biological sample" or "DNA contained in the

biological sample" will be understood to mean either the DNA present in the biological sample considered, or optionally the cDNA obtained after the action of a reverse transcriptase-type enzyme on the RNA present in the said biological sample.

Another aim of the present invention consists in a method according to the invention, characterized in that it comprises the following steps:

a) bringing a nucleotide probe according to the invention into contact with a biological sample, the DNA contained in the biological sample having, where appropriate, been previously made accessible to hybridization, under conditions allowing the hybridization of the probe to the DNA of a bacterium belonging to the species *Chlamydia trachomatis* or to an associated microorganism;

b) detecting the hybrid formed between the nucleotide probe and the DNA in the biological sample.

The present invention also relates to a method according to the invention, characterized in that it comprises the following steps:

a) bringing a nucleotide probe immobilized on a support according to the invention into contact with a biological sample, the DNA in the sample having, where appropriate, been previously made accessible to hybridization, under conditions allowing the hybridization of the probe to the DNA of a bacterium belonging to the species *Chlamydia trachomatis* or to an associated microorganism;

b) bringing the hybrid formed between the nucleotide probe immobilized on a support and the DNA contained in the biological sample, where appropriate after removal of the DNA in the biological sample which has not hybridized with the probe, into contact with a labelled nucleotide probe according to the invention;

c) detecting the new hybrid formed in step b).

According to an advantageous embodiment of the method for the detection and/or the identification defined above, it is characterized in that, prior to

step a), the DNA in the biological sample is amplified beforehand with the aid of at least one primer according to the invention.

The invention relates, in addition, to a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a nucleotide probe according to the invention;

b) where appropriate, the reagents necessary for carrying out a hybridization reaction;

c) where appropriate, at least one primer according to the invention as well as the reagents necessary for a DNA amplification reaction.

The invention also relates to a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a nucleotide probe, called capture probe, according to the invention;

b) an oligonucleotide probe, called detection probe, according to the invention;

c) where appropriate, at least one primer according to the invention as well as the reagents necessary for a DNA amplification reaction.

The invention also relates to a kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) at least one primer according to the invention;

b) where appropriate, the reagents necessary for carrying out a DNA amplification reaction;

c) where appropriate, a component which makes it possible to check the sequence of the amplified fragment, more particularly an oligonucleotide probe according to the invention.

The invention relates, in addition, to a kit or

set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, or for the detection and/or the identification of a microorganism characterized in that it comprises a DNA chip according to the invention.

The invention also relates to a method or to a kit or set according to the invention for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis*, characterized in that the said primer and/or the said probe according to the invention are chosen from the nucleotide sequences specific to the species *Chlamydia trachomatis*, in that the said polypeptides according to the invention are chosen from the polypeptides specific to the species *Chlamydia trachomatis* and in that the said antibodies according to the invention are chosen from the antibodies directed against the polypeptides according to the invention chosen from the polypeptides specific to the species *Chlamydia trachomatis*.

Preferably, the said method or the said kit or set above according to the invention, for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* is characterized in that the said primer and/or the said probe or the said polypeptides are chosen from the nucleotide sequences or polypeptides according to the invention which have been identified as being specific to the species *Chlamydia trachomatis* and in that the said antibodies according to the invention are chosen from the antibodies directed against the polypeptides according to the invention chosen from the polypeptides identified as being specific to the species *Chlamydia trachomatis*.

The invention relates, in addition, to a method or a kit or set according to the invention for the diagnosis of predispositions to, or of a condition caused by, genital diseases which are induced or worsened by a *Chlamydia trachomatis* infection.

The invention also relates to a method or a kit or set according to the invention for the diagnosis of predispositions to, or of conditions caused by, eye diseases induced or worsened by a *Chlamydia trachomatis* infection.

The invention also relates to a method or a kit or set according to the invention for the diagnosis of predispositions to, or of conditions caused by, systemic diseases, in particular of the lymphatic system, which are induced or worsened by a *Chlamydia trachomatis* infection.

According to another aspect, the subject of the invention is the use of polypeptides according to the invention, of cells transformed with a vector according to the invention and/or of transformed animals according to the invention, for the biosynthesis or the biodegradation of organic or inorganic compounds.

As has been mentioned above, the nucleotide sequences of the invention were identified by homology with sequences known to encode, for example, polypeptides or fragments of enzymatic polypeptides involved in the biosynthesis or the biodegradation of organic or inorganic molecules.

It is thus possible to use the said polypeptides of the invention in a similar manner for the biosynthesis or the biodegradation of organic or inorganic compounds of industrial or therapeutic interest (called compounds of interest).

Among these polypeptides, there may be mentioned in particular the enzymes involved in metabolism, such as the proteolytic enzymes, amino transferases, glucose metabolism, or the enzymes which may be used in the biosynthesis of sugars, amino acids, fatty acids, polypeptides, nucleotides, nucleic acids or any other organic or inorganic compound or in the biodegradation of organic or inorganic compounds.

Among these polypeptides, there may be mentioned, in addition, the mutated or modified enzymes corresponding to mutated or modified polypeptides

according to the invention which may also be used for the biosynthesis or the biodegradation of organic or inorganic compounds at the industrial level, such as, for example, the production of compounds of interest, the reprocessing of manufacturing residues applied to the food industries, to the papermaking industry or to the chemical and pharmaceutical industries.

The methods of biosynthesis or biodegradation of organic or inorganic compounds, characterized in that they use a polypeptide or one of its fragments according to the invention, transformed cells according to the invention and/or a transformed animal according to the invention, also form part of the invention.

The invention relates, in addition, to the use of a nucleotide sequence according to the invention, of a polypeptide according to the invention, of an antibody according to the invention, of a cell according to the invention, and/or of a transformed animal according to the invention, for the selection of an organic or inorganic compound capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or of modifying the cellular replication of eukaryotic or prokaryotic cells or capable of inducing, inhibiting or worsening the pathologies linked to an infection by *Chlamydia trachomatis* or one of its associated microorganisms.

The invention also comprises a method of selecting compounds capable of binding to a polypeptide or one of its fragments according to the invention, capable of binding to a nucleotide sequence according to the invention, or capable of recognizing an antibody according to the invention, and/or capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or of modifying the growth or the cellular replication of eukaryotic or prokaryotic cells, or capable of inducing, inhibiting or worsening, in an animal or human organism, the pathologies linked to an infection by *Chlamydia trachomatis* or one of its associated microorganisms, characterized in that it

comprises the following steps:

a) bringing the said compound into contact with the said polypeptide, the said nucleotide sequence, with a transformed cell according to the invention and/or administering the said compound to a transformed animal according to the invention;

b) determining the capacity of the said compound to bind with the said polypeptide or the said nucleotide sequence, or to modulate, regulate, induce or inhibit the expression of genes, or to modulate growth or cellular replication, or to induce, inhibit or worsen in the said transformed animal, the pathologies linked to an infection by *Chlamydia trachomatis* or one of its associated microorganisms.

The transformed cells and/or animals according to the invention may advantageously serve as a model and may be used in methods for studying, identifying and/or selecting compounds capable of being responsible for pathologies induced or worsened by *Chlamydia trachomatis,* or capable of preventing and/or of treating these pathologies such as, for example, genital, eye or systemic diseases, especially of the lymphatic system. In particular, the transformed host cells, in particular bacteria of the *Chlamydia* family whose transformation with a vector according to the invention may, for example, increase or inhibit its infectivity, or modulate the pathologies usually induced or worsened by the infection, may be used to infect animals in which the onset of pathologies will be monitored. These nontransformed animals, infected for example with transformed *Chlamydia* bacteria, may serve as a study model. In the same manner, the transformed animals according to the invention may, for example, exhibit predispositions to genital and/or eye and/or systemic diseases, especially of the lymphatic system, and thus be used in methods for selecting compounds capable of preventing and/or of treating the said diseases. The said methods using the said transformed cells and/or transformed animals form part

of the invention.

The compounds capable of being selected may be organic compounds such as polypeptides or carbohydrates or any other organic or inorganic compounds already known, or new organic compounds produced using molecular modelling techniques and obtained by chemical or biochemical synthesis, these techniques being known to persons skilled in the art.

The said selected compounds may be used to modulate the growth and/or the cellular replication of *Chlamydia trachomatis* or any other associated micro-organism and thus to control infection by these microorganisms. The said compounds according to the invention may also be used to modulate the growth and/or the cellular replication of all eukaryotic or prokaryotic cells, in particular tumour cells and infectious microorganisms, for which the said compounds will prove active, the methods which make it possible to determine the said modulations being well known to persons skilled in the art.

Compound capable of modulating the growth of a microorganism is understood to designate any compound which makes it possible to act, to modify, to limit and/or to reduce the development, the growth, the rate of proliferation and/or the viability of the said microorganism.

This modulation may be achieved, for example, by an agent capable of binding to a protein and thus of inhibiting or of potentiating its biological activity, or capable of binding to a membrane protein of the outer surface of a microorganism and of blocking the penetration of the said microorganism into the host cell or of promoting the action of the immune system of the infected organism directed against the said microorganism. This modulation may also be achieved by an agent capable of binding to a nucleotide sequence of a DNA or RNA of a microorganism and of blocking, for example, the expression of a polypeptide whose biological or structural activity is necessary for the

growth or for the reproduction of the said microorganism.

Associated microorganism is understood to designate in the present invention any microorganism whose gene expression may be modulated, regulated, induced or inhibited, or whose growth or cellular replication may also be modulated by a compound of the invention. Associated microorganism is also understood to designate in the present invention any microorganism containing nucleotide sequences or polypeptides according to the invention. These microorganisms may, in some cases, contain polypeptides or nucleotide sequences identical or homologous to those of the invention may also be detected and/or identified by the detection and/or identification methods or kit according to the invention and may also serve as a target for the compounds of the invention.

The invention relates to the compounds capable of being selected by a method of selection according to the invention.

The invention also relates to a pharmaceutical composition comprising a compound chosen from the following compounds:
a nucleotide sequence according to the invention;
a polypeptide according to the invention;
a vector according to the invention;
an antibody according to the invention; and
a compound capable of being selected by a method of selection according to the invention, optionally in combination with a pharmaceutically acceptable vehicle.

An effective quantity is understood to designate a sufficient quantity of the said compound or antibody, or of a polypeptide of the invention, which makes it possible to modulate the growth of *Chlamydia trachomatis* or of an associated microorganism.

The invention also relates to a pharmaceutical composition according to the invention for the prevention or the treatment of an infection by a bacterium belonging to the species *Chlamydia*

trachomatis or by an associated microorganism.

The invention relates, in addition, to an immunogenic and/or vaccine composition, characterized in that it comprises one or more polypeptides according to the invention and/or one or more hybrid polypeptides according to the invention.

The invention also comprises the use of a transformed cell according to the invention, for the preparation of a vaccine composition.

The invention also relates to a vaccine composition, characterized in that it contains a nucleotide sequence according to the invention, a vector according to the invention and/or a transformed cell according to the invention.

The invention also relates to the vaccine compositions according to the invention, for the prevention or the treatment of an infection by a bacterium belonging to the species *Chlamydia trachomatis* or by an associated microorganism.

Preferably, the immunogenic and/or vaccine compositions according to the invention intended for the prevention and/or the treatment of an infection by *Chlamydia trachomatis* or by an associated microorganism will be chosen from the immunogenic and/or vaccine compositions comprising a polypeptide or one of its fragments corresponding to a protein, or one of its fragments, of the cellular envelope of *Chlamydia trachomatis*. The vaccine compositions comprising nucleotide sequences will also preferably comprise nucleotide sequences encoding a polypeptide or one of its fragments corresponding to a protein, or one of its fragments, of the cellular envelope of *Chlamydia trachomatis*.

Among these preferred immunogenic and/or vaccine compositions, the most preferred are those comprising a polypeptide or one of its fragments, or a nucleotide sequence or one of its fragments whose sequences are chosen from the nucleotide or amino acid sequences identified in this functional group and

listed above.

The polypeptides of the invention or their fragments entering into the immunogenic compositions according to the invention may be selected by techniques known to persons skilled in the art, such as for example on the capacity of the said polypeptides to stimulate T cells, which results, for example, in their proliferation or the secretion of interleukins, and which leads to the production of antibodies directed against the said polypeptides.

In mice, in which a weight dose of the vaccine composition comparable to the dose used in humans is administered, the antibody reaction is tested by collecting serum followed by a study of the formation of a complex between the antibodies present in the serum and the antigen of the vaccine composition, according to the customary techniques.

According to the invention, the said vaccine compositions will be preferably in combination with a pharmaceutically acceptable vehicle and, where appropriate, with one or more appropriate immunity adjuvants.

Various types of vaccines are currently available for protecting humans against infectious diseases: attenuated live microorganisms (*M. bovis* - BCG for tuberculosis), inactivated microorganisms (influenza virus), acellular extracts (*Bordetella pertussis* for whooping cough), recombinant proteins (hepatitis B virus surface antigen), polysaccharides (pneumococci). Experiments are underway on vaccines prepared from synthetic peptides or from genetically modified microorganisms expressing heterologous antigens. Even more recently, recombinant plasmid DNAs carrying genes encoding protective antigens were proposed as an alternative vaccine strategy. This type of vaccination is carried out with a particular plasmid derived from an *E. coli* plasmid which does not replicate *in vivo* and which encodes only the vaccinal protein. Animals were immunized by simply injecting the

naked plasmid DNA into the muscle. This technique leads to the expression of the vaccine protein *in situ* and to a cell-type (CTL) and a humoral type (antibody) immune response. This double induction of the immune response is one of the main advantages of the technique of vaccination with naked DNA.

The vaccine compositions comprising nucleotide sequences or vectors into which the said sequences are inserted are in particular described in International Application No. WO 90/11092 and also in International Application No. WO 95/11307.

The nucleotide sequence constituting the vaccine composition according to the invention may be injected into the host after having been coupled to compounds which promote the penetration of this polynucleotide inside the cell or its transport up to the cell nucleus. The resulting conjugates may be encapsulated into polymeric microparticles, as described in International Application No. WO 94/27238 (Medisorb Technologies International).

According to another embodiment of the vaccine composition according to the invention, the nucleotide sequence, preferably a DNA, is complexed with the DEAE-dextran (Pagano et al., 1967) or with nuclear proteins (Kaneda et al., 1989), with lipids (Felgner et al., 1987) or encapsulated into liposomes (Fraley et al., 1980) or alternatively introduced in the form of a gel facilitating its transfection into the cells (Midoux et al., 1993, Pastore et al., 1994). The polynucleotide or the vector according to the invention may also be in suspension in a buffer solution or may be combined with liposomes.

Advantageously, such a vaccine will be prepared in accordance with the technique described by Tacson et al. or Huygen et al. in 1996 or alternatively in accordance with the technique described by Davis et al. in International Application No. WO 95/11307.

Such a vaccine may also be prepared in the form of a composition containing a vector according to the

invention, placed under the control of regulatory elements allowing its expression in humans or animals. It is possible, for example, to use, as vector for the *in vivo* expression of the polypeptide antigen of interest, the plasmid pcDNA3 or the plasmid pcDNA1/neo, both marketed by Invitrogen (R & D Systems, Abingdon, United Kingdom). It is also possible to use the plasmid V1Jns.tPA, described by Shiver et al. in 1995. Such a vaccine will advantageously comprise, in addition to the recombinant vector, a saline solution, for example a sodium chloride solution.

A pharmaceutically acceptable vehicle is understood to designate a compound or a combination of compounds entering into a pharmaceutical or vaccine composition which does not cause side effects and which makes it possible, for example, to facilitate the administration of the active compound, to increase its life and/or its efficacy in the body, to increase its solubility in solution or alternatively to enhance its preservation. These pharmaceutically acceptable vehicles are well known and will be adapted by persons skilled in the art according to the nature and the mode of administration of the active compound chosen.

As regards the vaccine formulations, these may comprise appropriate immunity adjuvants which are known to persons skilled in the art, such as, for example, aluminium hydroxide, a representative of the family of muramyl peptides such as one of the peptide derivatives of N-acetyl-muramyl, a bacterial lysate, or alternatively incomplete Freund's adjuvant.

Preferably, these compounds will be administered by the systemic route, in particular by the intravenous route, by the intramuscular, intradermal or subcutaneous route, or by the oral route. More preferably, the vaccine composition comprising polypeptides according to the invention will be administered several times, spread out over time, by the intradermal or subcutaneous route.

Their optimum modes of administration, dosages

and galenic forms may be determined according to criteria which are generally taken into account in establishing a treatment adapted to a patient, such as for example the patient's age or body weight, the seriousness of his general condition, tolerance of the treatment and the side effects observed.

The invention comprises the use of a composition according to the invention for the treatment or the prevention of genital diseases which are induced or worsened by *Chlamydia trachomatis*.

Finally, the invention comprises the use of a composition according to the invention for the treatment or the prevention of eye diseases which are induced or worsened by the presence of *Chlamydia trachomatis*.

Finally, the invention comprises the use of a composition according to the invention for the treatment or the prevention of systemic diseases, especially of the lymphatic system, which are induced or worsened by the presence of *Chlamydia trachomatis*.

Other characteristics and advantages of the invention appear in the following examples and figures:

Legend to the figures :

Figure 1 : Line for the production of *Chlamydia trachomatis* sequences

Figure 2 : Analysis of the sequences and assembling

Figure 3 : Finishing techniques

Figure 3a) : Assembly map

Figure 3b) : Determination and use of the orphan ends of the contigs

EXAMPLES

Experimental procedures

Cells

The *Chlamydia trachomatis* LGV2 strain used is identified to have over 98% homology with the outer membrane protein sequences omp1 (CHTMOMPA) and omp2 (CHTOMP2A) of the *Chlamydia trachomatis* serovar L2/434/Bu strain.

The *Chlamydia trachomatis* LGV2 strain is cultured on mouse fibroblasts (McCoy cells), obtained from the American Type Culture Collection, under the reference ATCC CRL-1696.

Culture of the cells

The mouse fibroblasts are cultured in 75-ml cell culture flasks (Corning). The culture medium is Dulbecco's modified cell culture medium (Gibco BRL No. 04101965) supplemented with MEM amino acids (Gibco BRL - No. 04301140) L (5 ml per 500 ml of medium) and 5% foetal calf serum (Gibco BRL No. 10270 batch 40G8260K) without antibiotics or antifungals.

The cell culture stock is maintained in the following manner. The cell cultures are examined under an inverted microscope. 24 hours after confluence, each cellular lawn is washed with PBS (Gibco BRL No. 04114190), rinsed and then placed for 5 min in an oven in the presence of 3 ml of trypsine (Gibco BRL No. 25200056). The cellular lawn is then detached and then resuspended in 120 ml of culture medium, the whole is stirred in order to make the cellular suspension homogeneous. 30 ml of this suspension are then distributed per cell culture flask. The flasks are kept in a CO_2 oven (5%) for 48 hours at a temperature of 37°C. The cell stock is maintained so as to have available daily 16 flasks of subconfluent cells. It is these subconfluent cells which will be used so as to be infected with Chlamydia. 25-ml cell culture flasks are

also used, these flasks are prepared in a similar manner but the volumes used for maintaining the cells are the following: 1 ml of trypsine, 28 ml of culture medium to resuspend the cells, 7 ml of culture medium are used per 25-ml flask.

Infection of the cells with Chlamydia

Initially, the Chlamydiae are obtained frozen (at -70°C), in suspension in a volume of 1 millilitre. This preparation is slowly thawed, 500 µl are collected and brought into contact with subconfluent cells, which are obtained as indicated above, in a 25-ml cell culture flask, containing 1 ml of medium, so as to cover the cells. The flask is then centrifuged at 2000 rpm in a "swing" rotor for microtitre plates, the centrifuge being maintained at a temperature of 35°C. After centrifugation, the two flasks are placed in an oven at 35°C for three hours. 6 ml of culture medium containing cycloheximide (1 µg/ml) are then added and the flask is stored at 35°C. After 48 hours, the level of infection is evaluated by direct immunofluorescence and by the cytopathogenic effect caused to the cells.

Direct immunofluorescence

Starting with infected cells, which were obtained as indicated above, a cellular smear is deposited with a Pasteur pipette on a microscope slide. The cellular smear is fixed with acetone for 10 minutes; after draining the acetone, the smear is covered with 30 µl of murine monoclonal antibodies directed against MOMP (major outer membrane protein) of Chlamydia (Syva, Biomérieux) labelled with fluorescein isothiocyanate. The whole is then incubated in a humid chamber at a temperature of 37°C. The slides are then rinsed with water, slightly dried, and then after depositing a drop of mounting medium, a coverslip is mounted before reading. The reading is carried out with the aid of a fluorescence microscope equipped with the required filters (excitation at 490 nm, emission at 520 nm).

Harvesting of the *Chlamydia trachomatis*

After checking the infection by direct immunofluorescence, carried out as indicated above, the culture flasks are opened under a sterile cabinet, sterile glass beads with a diameter of the order of a millimetre are placed in the flask. The flask is closed and then vigorously stirred while being maintained horizontally, the cellular lawn at the bottom, so that the glass beads can have a mechanical action on the cellular lawn. Most of the cells are thus detached or broken; the effect of the stirring is observed under an optical microscope so as to ensure proper release of Chlamydiae.

Large-scale infection of the cell cultures

The product of the Chlamydiae harvest (culture medium and cellular debris) is collected with a pipette, and distributed into three cell culture flasks containing subconfluent L cells, obtained as indicated above. The cells thus inoculated are placed under gentle stirring (swing) in an oven at 35°C. After one hour, the flasks are kept horizontally in an oven so that the culture medium covers the cells for 3 hours. 30 ml of culture medium containing actydione (1 µg/ml) are then added to each of the flasks. The culture flasks are then stored at 35°C for 48 hours. The cells thus infected are examined under an optical microscope after 24 hours, the cytopathogenic effect is evaluated by the appearance of cytoplasmic inclusions which are visible under an inverted optical microscope. After 48 hours, the vacuoles containing the Chlamydiae occupy the cytoplasm of the cell and push the cell nucleus sideways. At this stage, numerous cells are spontaneously destroyed and have left free elementary bodies in the culture medium. The Chlamydiae are harvested as described above and are either frozen at -80°C or used for another propagation.

Purification of the Chlamydiae

The product of the Chlamydia harvests, stored at -80°C, is thawed on a water bath at room

temperature. After thawing, each tube is vigorously stirred for one minute and immersed for one minute in an ultrasound tank (BRANSON 1200); the tubes are then stirred by inverting before being centrifuged for 5 min at 2000 rpm. The supernatant is carefully removed and kept at cold temperature (ice). The supernatant is vigorously stirred and then filtered on nylon filters having pores of 5 microns in diameter on a support (Nalgene) allowing a delicate vacuum to be established under the nylon filter. For each filtration, three nylon filters are superposed; these filters are replaced after every 40 ml of filtrate. Two hundred millilitres of filtration product are kept at cold temperature, and then after stirring by inverting, are centrifuged at 10,000 rpm for 90 min, the supernatant is removed and the pellet is taken up in 10 ml of 10 mM Tris, vigorously vortexed and then centrifuged at 10,000 rpm for 90 min. The supernatant is removed and the pellet is taken up in a buffer (20 mM Tris pH 8.0, 50 mM KCl, 5 mM $MgCl_2$) to which 800 units of DNAse I (Boehringer) are added. The whole is kept at 37°C for one hour. One ml of 0.5 M EDTA is then added, and the whole is vortexed and frozen at -20°C.

Preparation of the DNA

The Chalmydiae purified above are thawed and subjected to a proteinase K (Boehringer) digestion in a final volume of 10 ml. The digestion conditions are the following: 0.1 mg/ml proteinase K, 0.1 × SDS at 55°C, stirring every 10 min. The product of digestion is then subjected to a double extraction with phenol-chloroform, two volumes of ethanol are added and the DNA is directly recovered with a Pasteur pipette having one end in the form of a hook. The DNA is dried on the edge of the tube and then resuspended in 500 µl of 2 mM Tris pH 7.5. The DNA is stored at 4°C for at least 24 hours before being used for the cloning.

Cloning of the DNA

After precipitation, the DNA is quantified by measuring the optical density at 260 nm. Thirty µg of

Chlamydia DNA are distributed into 10 tubes of 1.5 ml and diluted in 300 µl of water. Each of the tubes is subjected to 10 applications of ultrasound lasting for 0.5 sec in a sonicator (Unisonix XL2020). The contents of the 10 tubes are then grouped and concentrated by successive extractions with butanol (Sigma B1888) in the following manner: two volumes of butanol are added to the dilute DNA mixture. After stirring, the whole is centrifuged for five minutes at 2500 rpm and the butanol is removed. This operation is repeated until the volume of the aqueous phase is less than 1 ml. The DNA is then precipitated in the presence of ethanol and of 0.5 M sodium acetate pH 5.4, and then centrifuged for thirty minutes at 15,000 rpm at cold temperature (4°C). The pellet is washed with 75% ethanol, centrifuged for five minutes at 15,000 rpm and dried at room temperature. A tenth of the preparation is analysed on a 0.8% agarose gel. Typically, the size of the DNA fragments thus prepared is between 200 and 8000 base pairs.

To allow the cloning of the DNA obtained, the ends are repaired. The DNA is distributed in an amount of 10 µg/tube, in the following reaction medium: 100 µl final volume, 1 × buffer (Biolabs 201L), 0.5 µl BSA 0.05 mg/ml, 0.1 mM dATP, 0.1 mM each of dGTP, dCTP or dTTP, 60,000 IU T4 DNA polymerase. The reaction is incubated for thirty minutes at 16°C. The contents of each of the tubes are then grouped before carrying out an extraction with phenol-chloroform and then precipitating the aqueous phase as described above. After this step, the DNA thus prepared is phosphorylated. For that, the DNA is distributed into tubes in an amount of 10 µg per tube, and then in a final volume of 50 µl, the reaction is prepared in the following manner: 1 mM ATP, 1 × kinase buffer, 10 IU T4 polynucleotide kinase (Biolabs 201L). The preparation is incubated for thirty minutes at 37°C. The contents of the tubes are combined and a phenol-chloroform extraction and then a precipitation are carried out in

order to precipitate the DNA. The latter is then suspended in 1 µl of water and then the DNA fragments are separated according to their size on a 0.8% agarose gel (1 × TAE). The DNA is subjected to an electric field of 5 V/cm and then visualized on a UV table. The fragments whose size varies between 1200 and 2000 base pairs are selected by cutting out the gel. The gel fragment thus isolated is placed in a tube and then the DNA is purified with the Qiaex kit (20021 Qiagen), according to the procedure provided by the manufacturer.

Preparation of the vector

14 µg of the cloning vector pGEM-5Zf (Proméga P2241) are diluted in a final volume of 150 µl and are subjected to digestion with the restriction enzyme EcoRV 300 IU (Biolabs 195S) according to the protocol and with the reagents provided by the manufacturer. The whole is placed at 37°C for 150 min and then distributed in the wells of a 0.8% agarose gel subjected to an electric field of 5 V/cm. The linearized vector is visualized on a UV table, isolated by cutting out the gel and then purified by the Qiaex kit (Qiagen 20021) according to the manufacturer's recommendations. The purification products are grouped in a tube, the volume is measured and then half the volume of phenol is added and the whole is vigorously stirred for 1 min. Half the volume of chloroform-isoamyl alcohol 24:1 is added and vigorously stirred for 1 min. The whole is centrifuged at 15,000 rpm for 5 min at 4°C, the aqueous phase is recovered and transferred into a tube. The DNA is precipitated in the presence of 0.3 M sodium acetate, pH 5.4 and 3 volumes of ethanol and placed at -20°C for 1 hour. The DNA is then centrifuged at 15,000 rpm for 30 min at 4°C, the supernatant is removed while preserving the pellet, washed twice with 70% ethanol. After drying at room temperature, the DNA is suspended in 25 µl of water.

Phosphorylation of the vector

25 µl of the vector prepared in the preceding

step are diluted in a final volume of 500 µl of the following reaction mixture:

After repair, the DNA is subjected to a phenol-chloroform extraction and a precipitation, the pellet is then taken up in 10 µl of water, the DNA is quantified by measuring the optical density at 260 nm. The quantified DNA is ligated into the vector PGEm-5Zf(+) prepared by the restriction enzyme EcoRV and dephosphorylated (see preparation of the vector). The ligation is carried out under three conditions which vary in the ratio between the number of vector molecules and the number of insert molecules. Typically, an equimolar ratio, a ratio of 1:3 and a ratio of 3:1 are used for the ligations which are, moreover, carried out under the following conditions: vector PGEm-5Zf(+) 25 ng, cut DNA, ligation buffer in a final volume of 20 µl with T4 DNA ligase (Amersham E70042X); the whole is then placed in a refrigerator overnight and then a phenol-chloroform extraction and a precipitation are carried out in a conventional manner. The pellet is taken up in 5 µl of water.

Transformation of the bacteria

Plating of the bacteria

Petri dishes containing LB Agar medium containing ampicillin (50 µg/ml), Xgal (280 µg/ml) [5-bromo-4-chloro-indolyl-beta-D-galactopyranoside (Sigma B-4252)], IPTG (140 µg/ml) [isopropyl-beta-D-thiogalactoside (Sigma I-6758)] are used, 50 and 100 µl of bacteria are plated for each of the ligations. The Petri dishes are placed upside down at 37°C for 15 to 16 hours in an oven. The number of "recombinant" positive clones is evaluated by counting the white colonies and the blue colonies which are thought to contain the vector alone.

Evaluation of the "recombinant" positive clones

Ninety-four white colonies and two blue colonies are collected with the aid of sterile cones and are deposited at the bottom of the wells of plates designed for carrying out the amplification techniques.

30 µl of the following reaction mixture are added to each well: 1.7 mM $MgCl_2$, 0.2 mM each of dATP, dCTP, dGTP and dTTP, two synthetic oligonucleotides corresponding to sequences flanking the cloning site on either side and orienting the synthesis of the DNA in a convergent manner (0.5 µM RP and PU primers, 1 U TAQ polymerase (GibcoBRL 18038-026)).

The colonies thus prepared are subjected to a temperature of 94°C for 5 min and then to 30 thermal cycles composed of the following steps: 94°C for 40 s, 50°C for 30 s, 72°C for 180 s. The reaction is then kept for 7 min at 72°C and then kept at 4°C.

The amplification products are deposited on an agarose gel (0.8%), stained with ethidium bromide, subjected to electrophoresis, and then analysed on an ultraviolet table. The presence of an amplification fragment having a size greater than 500 base pairs indicates the presence of an insert. The bacterial clones are then prepared so as to study the sequence of their insert.

Sequencing

To sequence the inserts of the clones obtained as above, these were amplified by PCR on bacteria cultures carried out overnight using the primers for the vectors flanking the inserts. The sequence of the ends of these inserts (on average 500 bases on each side) was determined by automated fluorescent sequencing on an ABI 377 sequencer, equipped with the ABI Prism DNA Sequencing Analysis software (version 2.1.2).

Analysis of the sequences

The sequences obtained by sequencing in a high-yield line (Figure 1) are stored in a database; this part of the production is independent of any treatment of the sequences. The sequences are extracted from the database, avoiding all the regions of inadequate quality, that is to say the regions for which uncertainties are observed on the sequence at more than 95%. After extraction, the sequences are introduced into a processing line, the diagram of which is

described in Figure 2. In a first path of this processing line, the sequences are assembled by the Gap4 software from R. Staden (Bonfield et al., 1995) (OS UNIX/SUN Solaris); the results obtained by this software are kept in the form of two files which will be used for a subsequent processing. The first of these files provides information on the sequence of each of the contigs obtained. The second file represents all the clones participating in the composition of all the contigs as well as their positions on the respective contigs.

The second processing path uses a sequence assembler (TIGR-Asmg assembler UNIX/SUN Solaris); the results of this second processing path are kept in the form of a file in the TIGR-Asmg format which provides information on the relationship existing between the sequences selected for the assembly. This assembler is sometimes incapable of linking contigs whose ends overlap over several hundreds of base pairs.

The results obtained from these two assemblers are compared with the aid of the BLAST program, each of the contigs derived from one assembly path being compared with the contigs derived from the other path.

For the two processing paths, the strict assembly parameters are fixed (95% homology, 30 superposition nucleotides). These parameters avoid 3 to 5% of the clones derived from eukaryotic cells being confused with sequences obtained from the clones derived from *Chlamydia trachomatis*. The eukaryotic sequences are however preserved during the course of this project; the strategy introduced, which is described below, will be designed, inter alia, not to be impeded by these sequences derived from contaminating clones.

The results of these two assemblers are processed in a software developed for this project. This software operates on a Windows NT platform and receives, as data, the results derived from the STADEN software and/or the results derived from the TIGR-Asmg

assembler, the software, results, after processing of the data, in the determination of an assembly map which gives the proximity relationship and the orientation of the contigs in relation to one another (Figure 3a). Using this assembly map, the software determines all the primers necessary for finishing the project. This treatment, which will be detailed below, has the advantage of distinguishing the isolated sequences derived from the contaminations, by the DNA eukaryotic cells, of the small-sized sequences clearly integrated into the project by the relationships which they establish with contigs. In order to allow, without any risk of error, the arrangement and the orientation of the contigs in relation to one another, a statistical evaluation of the accuracy of the names "naming" of sequence is made from the results of "contigation". This evaluation makes it possible to give each of the clone plates, as well as each of the subsets of plates, a weight which is inversely proportional to probable error rate existing in the "naming" of the sequences obtained from this plate or from a subset of this plate. In spite of a low error rate, errors may occur throughout the steps of production of the clones and of the sequences. These steps are numerous, repetitive and although most of them are automated, others, like the deposition in the sequencers, are manual; it is then possible for the operator to make mistakes such as the inversion of two sequences. This type of error has a repercussion on the subsequent processing of the data, by resulting in relationships (between the contigs) which do not exist in reality, then in attempts at directed sequencing between the contigs which will end in failure. It is because of this that the evaluation of the naming errors is of particular importance since it allows the establishment of a probabilistic assembly map from which it becomes possible to determine all the clones which will serve as template to obtain sequences separating two adjacent contigs. Table 2 below gives the clones and the sequences of the primers initially

used during the initial operations.

Table 2: Directed sequencing

List of the names of clones and of the primers used to obtain the sequence of the regions separating two consensus sequences. The sequence was obtained using, as template, DNA preparations derived from these clones. The primer T7 is used systematically to serve for the sequencing of the control clones.

61-003-1-C3	GAAGATCCATGAGTCGATGC
61-003-1-D4	CCGCGAGCAATTAGATGACG
61-003-1-G6	AGTCGTCTGGACCTGTCCAG
61-003-1-H1	CGTGCTTTGAATAACGGGGA
61-003-2-G2	CACAGAGGGACTAAATAGGG
61-003-3-F11	CTGTATAGCCGCAGGGATTA
61-003-4-A9	GGGTAACTAGAACATTGCGT
61-003-4-B2	GTTGCACATCAAGCTGAGCA
61-003-4-B7	GGGAGAACGTAAAGCAAAGG
61-003-4-D10	GCTCAAATGCGAGAGGGAAA
61-003-4-G8	GGAAGAAGGAGGAAATACAG
61-004-1-B2	GGGGGGTCTTTCTAAAAGCC
61-004-1-B3	GGCGAACGGGAAAATTTCTC
61-004-2-E3	TCCTGGAGACCCTATCATTC
61-004-2-E5	GGGCTGGGGCTTTTAGTAAA
61-004-3-B9	GGGTATTAACCGACCACAGT
61-004-3-B10	GGCTTCACCTCGTGATAAAC
61-005-1-A5	TGCACCGGGCGAAGGAATGT
61-005-1-D11	GTGAGCCCACAACTTTCACA
61-005-1-E2	CGGATGATCGCTCAGCTAAT
61-005-2-D4	CCCTTCTTCGTTTTGGTAGC
61-005-2-D6	AACCCGTGCTGCTTGAGAAT
61-005-2-E7	GCTGTTGTGCTCCTTGTTCA
61-005-2-H3	GGAGAGAGAGTTAAGGTTTC
61-005-3-A4	GATCTCGTCCCTAGAACTTC
61-005-3-B4	GCCATTATTATCTGGGCGCT
61-005-3-C10	GCGGGCAATACAAGAACAAG
61-005-3-E1	CCCTGCCGTGGTGCTATATG
61-005-4-E2	AGCTTTGAGTCGGGGGAATG

61-006-1-B3	CCATGCTCGCTAATCGCTTA
61-006-2-B1	GAACTTTAGATGTGCACCTG
61-006-2-C4	GTTCCCTTCACAGGTGGATT
61-006-2-C11	CTGTAGAGAAAGGGCTTGCT
61-006-2-D5	GCGAGGGCGCAGATAAAAAA
61-006-2-E3	GTAGAGGCAGTTTCTAAGCG
61-006-3-A2	AGCTGCAGGACCAGCTCCAG
61-006-3-B6	CACTCTTTTGGTTCCTTGCG
61-006-3-B12	GTTCCGTTCGGGTCATAAAC
61-006-3-C3	CGAATCTCGAGCAAGTTTCC
61-006-3-H6	AGCGGAAGAAGCGCGCTTGT
61-007-1-G11	GTGATACGCGACGTGAGAAA
61-008-1-C9	CCCCATATACACGGACACTT
61-008-1-D6	GAAGTCTGCCGAGCACTTTT
61-008-1-D11	GGGACTAGTTCGACTCCAAT
61-008-1-F1	GGTGCTTCTGATCAGGTCAT
61-008-3-B5	CCGGGAATTCTCCTTGGTTT
61-008-4-A9	CGAGAACCACTTGTTTACTC
61-008-4-D5	GGGGTTAAGGATAAAGCCAC
61-008-4-F1	CCGTATTACTGTCTCGTTCC
61-009-1-D3	CATAACCGTTGCATGTTGGC
61-009-1-F4	GCGCGGAAGATTTCCAAACT
61-009-1-G11	ACCCAACACCCTTAGGCATA
61-009-2-B7	CCAACCCGAAGAATCTCAAG
61-009-2-H9	GCTTACGATAGCCAAGCGGT
61-009-3-C7	CCCTATAAGAGCTGGTCTCA
61-009-3-E8	ACATCACCCTGCTGAAGCTG
61-009-3-G1	CGGGAAACCTTACAGGCATT
61-009-3-G3	TCTCGCAAACGCGCTAGAAT
61-009-3-G12	ACTTCTTCGCCACTGATCGC
61-010-1-E10	GCGAATCCCTTCCCTTTTTC
61-010-1-F2	GCGATCTCACTTTCCCACAA
61-010-2-D1	GGAAGGGATCGTTCCCAGAA
61-010-3-D11	CTACTTGGCGATAGGCCTAA
61-010-3-E9	GCAGCAGTCTTCATGACGTT
61-011-1-A1	GCTCTTTTCGGAGTAGGAGC
61-011-1-D11	GGGGAGAGTTTGTTACGAAC
61-011-1-F1	CGATTACCAAGATCGGCAAG
61-011-1-H2	GGGTGTATCATTGATCTCGC

61-011-2-A11	GCGATTCGAGAAGGGAAGTT
61-011-2-B10	CGATCCAACTGGTCGTACAC
61-011-2-H2	TGGCATCCAATGCCTCTTGC
61-011-3-B1	GCGGTACTTTTCTTAGCGAA
61-011-4-A7	CCCGTTTCTTAAGGTTATCG
61-011-4-A8	GCCGTGACTACAAAAGGAAG
61-011-4-B3	CCGAAAGGATCGCTCGATTA
61-011-4-G8	CGCGATGTTATTAAGCACCC
61-012-1-F6	TTTACGCTACTCAGAGCCTC
61-012-2-A7	CCGAAGAAATATAAGCGCGG
61-012-2-H3	GACATCCCTACACCTTTAGC
61-012-4-A7	CCGCTGTGGATTTTTAGGAC
61-012-4-F4	GATGTCGGCACCTTCAGAAA
61-012-4-H5	CCCAAGTTTCTACTAGGAGC
61-013-1-B7	GTTTCTAGCGTTCGTCTTCG
61-013-1-E1	GCTACTTGATAGATGCCTGG
61-013-1-F9	GCGTTAATTGCATGCGGAGG
61-014-2-E11	CGGAACAACCCCCTAAAGAA
61-014-2-F1	CTACAGCTGGAGGGATAGTT
61-014-2-G9	CTTCCCCAACTCTTTTCCAG
61-014-2-H6	TAGCGCTCATGGGACTGATT
61-014-3-F4	CGCTCATGGATGTAACGAAC
61-015-1-B2	CGCTCACCTTCTCGGTTAAT
61-015-1-B3	CTAGACAAACGAGCTCTTGG
61-015-2-B5	CCAAGTTGTTGCTCTGTACG
61-015-3-A1	CCCTCCCTCTTCCTTAGTTA
61-015-3-D7	CTCAATTGGCAGAGCATTCG
61-015-3-D8	CGCATTGCGGGTACAAAAAT
61-015-3-G5	GGCCTCTAAGTTTGGTTCTC
61-016-1-B4	GACGCTCTTGTTGGACTTCT
61-016-2-A1	GCTTCTTCCACAGATTCTCC
61-016-2-D12	CTCGAGCTTTTTCATCGTCC
61-016-3-B2	TACAGGGCATCGTGTAGAGG
61-016-3-B10	GGTGGGTTAACTCATCGTAG
61-016-4-B1	GGCTTTCCCCAAAAGAAGAG
61-016-4-C3	GCAAGGTCAATGCTTAGGAG
61-016-4-D3	CCGTCTGTGACTCCTATACT
61-017-1-E12	CTGCAACAGTCATCTCCTCT
61-017-2-C1	CCCAACTTGTGTCCAACCAT

	61-017-2-D7	GCGCTCCATGATTTTTGCTG
	61-017-2-G1	GGTCTTGCGACACCCTGTTT
	61-017-3-A7	GCTTGTCATTGGTAATGCGG
	61-017-3-B5	CCTGGAGAGCACAGAACAAT
5	61-017-3-H10	GGAAACTCCTTGCGACAAAG
	61-017-4-B2	CAAGGGCTCTCTCTGGCTAA
	61-017-4-E7	CACGTTGACCATGGTAAGAC
	61-017-4-G8	GATGCGATTACACGATCCCA
	61-017-4-H3	GCTAACCAAGCGGTTACCTT
10	61-018-1-B1	ATGATATTCACGCCTGCAGC
	61-018-2-A6	TACGGCTCTGGCGCTAGTGC
	61-018-2-G8	GGGGAATTTCTCTCACACCT
	61-018-3-C7	GCACAATCCTTAGCTCCAGA
	61-018-3-C8	ACCTTCTTTGAACCTCCGAG
15	61-018-4-D5	GCCTTATTTATTGGGGCAGG
	61-018-4-F1	ACGAACTACCTGCAAAGAGC
	61-018-4-H5	CACATACGTTTGAGTCGTGG
	61-019-3-C4	GAGAAGAGCTTCTCGTGAAG
	61-019-3-D2	TTTCGCTAAATCGCCCACGG
20	61-019-3-F1	GGACTGCTCTGTTTACCAGA
	61-019-3-G7	GCATACCGTAGGGCGCCTTC
	61-019-4-A3	CTTATCCATCGGGAGTATCC
	61-019-4-H6	TGGTGGATTATCTCCCCTCC
	61-020-1-H7	AATGCTGCGGTTGCTTCTTG
25	61-020-2-B9	TGATGGGCATGCAGTGCTCG
	61-020-3-B8	GCGCAACAATCTGCTATTTC
	61-020-3-D8	CGTAGTTCCTGCCTACTTGT
	61-020-3-E5	TCGGGTACGCGCTATTCTAT
	61-020-3-E6	GACCAACTAATTTAGCCAGC
30	61-020-4-B6	CTGCTAAGGACTTTGACGAC
	61-020-4-G6	GGTGAAGAGTTCTCTCCTGA
	61-021-1-A6	GCGGCATCAGCTACAAACAT
	61-021-1-E3	TCGTTTGACTCAACCGCAAG
	61-021-1-G12	CACGATCAACTTGCCTTCTG
35	61-021-2-H3	ATAGTTTGCACTGGGAGGAG
	61-021-4-A6	ACTCGCTTGATCGATGAGAG
	61-021-4-G7	GCTCGGATGTTCCTAATGCC
	61-021-4-H5	TTCCTCTCCAACCCTCATGT
	61-022-2-A7	CCTGTTGAGGGAGAGAAATC

61-022-2-C1	CAAGTAGCCAGTGACTCGCT
61-022-2-H11	CCCACTCTACCTGTTTACGA
61-022-3-A1	GACAAAAGGCCGAAAGAAGC
61-023-2-B8	CCGCCATAATCTGAGTTGGA
61-023-2-F9	GTTGGAGATCGTTATGGGAG
61-023-3-A12	GGCGATCCTACCCAAATAGA
61-023-3-E10	GATATGGCGATGTTCCTGCT
61-023-4-A9	GTTCCTACAGCAGTAGAAGC
61-024-1-B9	GAGAGGATGACGCTACTTGA
61-024-1-D11	GGAGAGCGAGTGACCTTATT
61-024-1-E12	GGGATGTTGCGAGAAGACAT
61-024-2-D4	CCATTCTTAGCTGCAGCGTA
61-024-3-A5	GCAATAGCTAAAGGGAAGGC
61-024-4-A10	GAGATACAGCAGAGAGGTCT
61-024-4-G5	CCTCGACGATACCTTGAGCC
61-025-1-D8	GCCATTTGAGACGATAAACC
61-025-3-D11	ATCTACACGCTCAGGCTATC
61-025-4-B12	CCGTGTGCCCGAAGATAATA
61-025-4-D8	CGTAATCCGGCAGTCCCAAA
61-026-3-C1	CAACCCTTCTGGAGAAATGC
61-026-4-E1	GCCTATCTTAGATCAGCTTC
61-026-4-H4	CGCTCCCTATAGTGAGTGAT
61-027-3-D4	GCTGCCTCCTGCGGCTCTTC
61-027-4-A1	GGGTCACCCGTATTTAGGTA
61-027-4-B3	GCTTAGGGAAAGATCCCCAA
61-027-4-C11	GCGAGGCTATCCGTTCTTTA
61-027-4-F9	CCTGACCTTTCTTATGTGCC
61-028-1-F5	TTGCGAGCCTCTCGTCCATC
61-028-1-F6	ATCACGACCGAACATCGCTT
61-028-2-A3	ATGGCTTAGGGCGTTCCGGC
61-028-2-A4	CAAGCTTCTCAAAGGTATCC
61-028-2-D11	TGATCTGGCCTGCTATCGTG
61-028-3-H10	GGCAGAGCTTCGCAATCATT
61-029-3-H8	GGGCTGAAAGTGGTTTCACA
61-031-2-A10	CAGGGGCTGTCATTCAAGAT
61-031-2-C11	CCTAGCAGACTTGTAAACCC
61-031-2-D9	GAAGATGGAGCTATGCAAGG
61-031-2-F6	AGGAAGAGAGTACCCTTGGA
70-001-1-A2	AGCACAGGCCCGGTTAATAA

70-001-1-B10	TCCCACAGACATGCCCAGAG
70-001-1-C8	CCATGGATAGGTTAGGCTGA
70-001-1-E11	GCGTCCATCAGGCGTAGAAC
70-001-1-F5	CCCGTCTGGGTATACACAAT
70-001-1-G11	GGATTCTTACGGATAGCAAG
70-001-1-H4	CCGCGATTGTATCTTTGGCT
70-001-2-D3	ACAAGCAGCAAAATTCCCTG
70-001-2-D7	CACAACTGCTGTAGCTGAGA
70-001-2-H3	CGGAAACGCTGTTTCGTACT
70-002-2-A6	GGAAAGAACGGAGGCTTCTT
70-002-2-E10	ACCGACGAGTTCGCTCTGTT
70-002-4-D5	GTTTGGAAGCAACGATCTGC
70-002-4-E2	ATTCGCTTGCAGAGGGCCTC
70-002-4-E11	GGGGCTGTTCCAGGAAGGTT
70-002-4-G11	CGAACCCCTCTGAAAAACGA
70-003-1-F6	GCGTCCAAAAAGCGTGTCAA
70-003-2-D4	CAGAAGCATAGCTATGTCGT
70-003-2-E5	CTCCATCAGCAACCTTATCC
70-003-2-F1	GCCGCTTCTGAAGATATTGG
70-003-2-G7	CTTCGGCAGAAGCTGCAGAG
70-003-2-H2	GAACTTCGGCTTGTAAAGCC
70-003-3-C7	CGCTTGGTGTATCAATCGTC
70-003-3-D8	AGTTCTGCTTCCTCCTTTGC
70-003-3-E3	CCAGCGGAAAGTATCTAGCA
70-003-3-E4	CACTGGAAGCCTCTACTCTA
70-003-3-F6	AAGGATCGCTATCGTAAGGC
70-003-4-A4	CCTTCTTGGGCTGCGGATAG
70-003-4-A11	GCTAGGCTTGCTTACGTTCT
70-003-4-B4	GCGCTTCCTTCTTCTAGAGA
70-003-4-C10	GCAACTTCCCAGTCTGGATG
70-003-4-D2	CTAACGGTCGGAGATCTTCC
70-003-4-D8	GCAAAACAGGGACTTCCTCT
70-003-4-E5	CTCGTAAGGGGTGCCGCACA
70-003-4-F7	GAGAAATTCCAAGCGTTCGC
70-003-4-H2	CCGATCTGTTATTCAGCGCA
70-003-4-H9	GCTTTATCCGTACGAGCAAC
70-004-2-D10	CCACAGCCATAATAAGCCGC
70-004-2-F1	GCATCGACACCATCACCATC
70-004-2-H8	GAGCATAGCGCTGTCGAAAT

70-004-3-B12	GCGAGGGAATGTTTGCTTTC
70-004-3-C10	CGGACTATGGTTTTCTGTCC
70-004-3-D4	GCCTGAAGAGTTACGAGAGC
70-004-3-D6	CCCTTCCTTCTCCCAGAATA
70-004-3-E10	GGAGGATCCGAAAAAACGCT
70-004-3-F7	CTAATAAGGCTCCGATCGGT
70-004-3-H7	CAGCCTCTAATGCCGAAGTA
70-004-4-A7	ACATTCGCCAATCGAGCGGG
70-004-4-A8	CGCTATGTTGACACTAGGCT
70-004-4-E1	AGACTAGGAAAAATGGGGCG
70-004-4-E6	GCTTCATCGATGGGAAGGTT
70-004-4-F8	CGCCTTCCTTCGTGGATTAT
70-004-4-F11	ACCAACCCCGGCTCCTGCAA
70-004-4-G6	CGACGGTTCTATTGTCCGAA
70-005-2-B1	ACACCGCTTCTTGTGTTGGA
70-005-4-D10	AATCGGATACCTAGCGCAAG
70-005-4-E7	CCTTGATAGCAGTCGTCATC
70-005-4-E8	GGGGCCAGTACATATTACTG
70-005-4-F3	CTGCAGGAATTTTGTCGTGG
88-001-2-A9	GCGGCATTTTACTTTGGAGC
88-001-2-E6	CGCCAAAGATTTCTTCGACG
88-001-3-E12	GTGATGGGACTTCAGCAAAC
88-001-3-F9	AGGGGTAGAGTCTGTGCTTA
88-001-3-G2	GACAGGGTTTGTCTGATCCA
88-001-3-G11	GGTTCGTTCCCATTTCGAGA
88-002-2-B5	GACATGCACGATGTAGCCAA
88-002-2-B6	AGATAAGCGCCGTGGATCTC
88-002-2-D2	CGAGAGAGAGAAGCTGTCAT
88-002-2-D10	GAGGAATAGAAGCCCTCCTT
88-002-2-G1	GGACTCTTACACGAACGAGA
88-002-3-D1	TACAGGGAGTTTGTCCTTGG
88-002-3-D12	CTCGATCCAGCATCTCGGTA
88-002-3-F3	GGCTCCTTCTCAACAGAAGC
88-002-3-F11	CCCCTTACCCATTCTCTACT
88-002-4-D8	CGCACAGAGCTATAGCATCT
88-003-2-B1	GAAGAGAGGAGAGTGTTTCC
88-003-2-B11	GGAAGCGTTAGAAGCTTTGG
88-003-2-C1	CAGAACTCCCTCCATCAAAG
88-003-2-D9	GGGATGGCGATGTTTAGAAG

	88-003-2-E9	CAGTTCGTACAGATACCGTC
	88-003-2-F11	CCTGGGATGCAACCACAAAA
	88-003-3-A1	CTGACAATGCCATGACAGGA
	88-003-3-A7	CCAAACCACCCGTTGCAATA
5	88-003-3-D8	CCATAGGATGGGCAGTCAAA
	88-003-3-G5	TCGGAAGAAAGAATCGCTTC
	T7	TAATACGACTCACTATAGGG
	610031D4_A	TCTTCCTGGCATCCGATCTG
	610033E8_A	GCACCGGCGATATAGAAGTT
10	610041D3_A	AAGCAGGATAACGGCGACAA
	610063B12_A	GAAACTTTGCGCAGGATTCC
	610081B5_A	TTTTACTGCCTCTGTATCGC
	610081G12_A	GGTGTAATCCCATTAGTGAC
	610083A3_A	CGGGATTTGCAAACTGCACA
15	610084A9_A	CGAGAACCACTTGTTTACTC
	610111F6_A	CTCCCGAGGTGATTTAAAGG
	610112B6_A	GACTGTCTGCAAAAGGCTCT
	610131B7_A	GTTTCTAGCGTTCGTCTTCG
	610141E2_A	CTTCCCTTCATAGATCAGCG
20	610142F1_A	CTACAGCTGGAGGGATAGTT
	610144G4_A	CAAGCAGCCATGAGACTATC
	610153D7_A	CCTTACGGATACGTTGGTTC
	610161G5_A	CTGTAGGGAGGTTATAGAGC
	610172G1_A	TACAAGAGGTCTTGCGACAC
25	610173H10_A	GAGGAAGCGCTTTTCTTTGC
	610182D6_A	GCCTCTGGAATCTTCCTTGT
	610194C5_A	CATGGCCAATCCGTATGGAT
	610201G3_A	GAGTTCCTTCAGAAGAGGCA
	610212H3_A	GAGGAGGGAGAGCATTATCA
30	610213E12_A	CACAGAGAAGTTTCTCCCCT
	610222F9_A	CTTTGCTCTGCTATGGGGAT
	610223A1_A	CTAGCATTGTAGCTGCCAGA
	610244E12_A	CCAAACAGTACTTGCGTCTC
	610244G5_A	GAGATGGGAGTTTCTGAACC
35	610254G11_A	TCATCTCTGCGGAGCGAGAG
	610264F7_A	CGTAGTGAAGCAGCGAGTCT
	610281A10_A	GAGCAGGAACGGGAGGAGAT
	610282D11_A	TGATCTGGCCTGCTATCGTG
	700011A9_A	CCGACCTATTGCTCCCAAAG

700012D7_A	CACAACTGCTGTAGCTGAGA
700012H3_A	CTCGGAAACGCTGTTTCGTA
700022C9_A	CCAGACTCCTTTGCAGAGTT
700031D10_A	GTAAGCCTTGCGCATTCTTC
700032C3_A	CGAGGTTTCTTATTCCCATC
700033D5_A	CGCAAGAATCACAAGTGTCG
700033H11_A	GGTAGCTGCGTTGGTAAATG
700034G11_A	CAGTAGACGATACACCAGAC
700042G1_A	GTAGGGGATCAGCGGTTAAA
700044E6_A	CCTGCAGAGCAATACAGTGT
700044H9_A	GAGCAGCATACCGTAGGGCG
700052B3_A	CTAAATCTGACATGCCGTGC
880013A8_A	GCGTAGATTTACTTCCGGAG
880013H9_A	CACTTCTGTGGGACTCCCTA
880014G12_A	GCTTTGTGGAAACATCGAGC
880014G6_A	GCGCACCACCCCTATTTTTT
880022D2_A	GCACTTCGTTTTCCAGTGCT
880022F1_A	CGTTCCATGCTTCAGCTGTT
880023H10_A	TCGGAATCGGATACCTAGCG
880024H3_A	CGGAGGATGATCCAGAATCG
880032A9_A	GGACCCTTATCTAACGTGAC
880032B5_A	GGTGGAACAACTTCCCTGTT
880032D5_A	GATGTCGGCACCTTCAGAAA
880032F11_A	CTTCTAAGAAGTTGGTCGGG
880033E1_A	TCAGGCCTTCCTACTTGAGC
880033E4_A	CCTCCAAACAACCATCCTGA
T7	TAATACGACTCACTATAGGG
610031D4_B	GAGCTGGAAAGGGCTGAAGT
610033E8_B	CCGCGAGCGAAACTTTTAAC
610041D3_B	TTCCTTTCCAATGGCCGTCA
610063B12_B	GTCTCCCCGCATAGTTTTTC
610081B5_B	GCGCTCAAAGTACAAGTCGG
610081G12_B	CCTGGGATGCAACCACAAAA
610083A3_B	CGATCCTACTTCCGTAGATG
610084A9_B	GAAGTCCCTCCTACTAAGGA
610111F6_B	GCTTTTGGTCCTTTCTCTGT
610112B6_B	GGACAGCAGCTCGAAGAAAT
610131B7_B	CAAGAAGCCTTGAGCCTTCA
610141E2_B	CGCTCTGGCAGTTTTTACTC

	610142F1_B	CCATCCTAATGCTCCCCATA
	610144G4_B	GTGCTCACTATGCTTTGTGG
	610153D7_B	CCTGTAGTCCCTGAGAGATT
	610161G5_B	GGAGATCACCTAACCTAGGA
5	610172G1_B	AGCCCAAATAGCCAGAAAGC
	610173H10_B	ACGAAACGATCCGAAACGAC
	610182D6_B	CTATGGGAGATATCCCAGGC
	610194C5_B	GGCGCTCTTCATGAACGATA
	610201G3_B	GGAAACTCAGACAGTAACGA
10	610212H3_B	GGAGTGGTTCTTGTCAGTAG
	610213E12_B	GCAGCCATATCACGAATTCC
	610222F9_B	GGGCAGGAGAAATCCCATCT
	610223A1_B	CGGTTTTCTCTCTCGTCAGC
	610244E12_B	GCGATTGCGATGATGAAGAC
15	610244G5_B	GCAAGCTTTTTTCCTCGACG
	610254G11_B	CCACTCTTTCCTATCCCAGA
	610264F7_B	AGGAGGAGTTAACGCTGGAT
	610281A10_B	ACTTCGCCTGATGGAGAGCG
	610282D11_B	CTATTCATGACTGCGTTGCC
20	700011A9_B	AAACACCCTCTCCTACGAAG
	700012D7_B	CCCAAGCGGAATCTTGTGAA
	700012H3_B	ACGGATCCGTCCAAGGAAAC
	700022C9_B	GCAACCTTCTCCTCCAATGG
	700031D10_B	GCGACAATTCTAACAGGAGG
25	700032C3_B	GTTGAGTAAGAGGAGAGTCT
	700033D5_B	AGCACTTCCTCAAGAAGTGC
	700033H11_B	GTTGGATGCTGTGCCTATTC
	700034G11_B	GGAAATCGGTGACGGAAGTT
	700042G1_B	GTTTGGAAGGTGAGGCTCTA
30	700044E6_B	TGTATGAAGTTGCTTGGGGC
	700044H9_B	TGATTCGGTAATCACGCCTC
	700052B3_B	CGCCTTCCTTCGTGGATTAT
	880013A8_B	GGCAATCCGACCTCATCTAA
	880013H9_B	CAAGACCCCTATCCTGACAA
35	880014G12_B	GGCTGGATCTAATGTGTCGT
	880014G6_B	GGCAGTACGGCAAGTAATGA
	880022D2_B	CGAGAGAGAGAAGCTGTCAT
	880022F1_B	GGACGTACTGCGTCTCTAAA
	880023H10_B	AATCGGCGACGTTCTCCGCC

	880024H3_B	CGAGGTTTCTTATTCCCATC
	880032A9_B	GTCCCGAAGATTGAAAGAGG
	880032B5_B	GTGAATGAGCAGACAAAGCG
	880032D5_B	GCTCCGCTTTCTGCATTGGT
5	880032F11_B	CCTGGGATGCAACCACAAAA
	880033E1_B	CGGATCTCGCATCAGCAAAA
	880033E4_B	CAAGCTTCGGTTCACAAAGC
	T7	TAATACGACTCACTATAGGG
	610082A6_B	GGATCGGCTTTCAATTCCAG
10	880032C1_B	GGTGTCAAGGAGCTAAACAC
	880013G2_B	GCAAGCTTCCCGATTTGAAG
	610164E5_B	ATCAATCCTCAGGATCTCGC
	700054A3_B	CTGACATCTTCCCCTACCGC
	880024F8_B	GCCATTATTATCTGGGCGCT
15	700033E1_B	CTGCCGCTAGCGAATTTGAT
	880013C11_B	GAGAAAATGCAGCGATTCCC
	610031E12_B	CTCCCCCCATGAAAAGCAAT
	610182B4_B	TGCTGTTGCAACAGGAATCC
	960050B8_B	GCAGGCAGGTTTTGTATGAC
20	610051A5_B	GCCTACACTATTTTGCACCG
	960060E3_B	CCTAAGCGAGCTCAAATGGA
	610173C2_B	CCAAGAAGCTTTTTCCAGCC
	610252B12_B	GAAGTCTGCCGAGCACTTTT
	610292H4_B	GCTATGTCCTTCATCCCAGA
25	880024E11_B	GCGCGGAAGATTTCCAAACT
	880033E8_B	ACCCAACACCCTTAGGCATA
	700042A6_B	GCTCCGCTTTCTGCATTGGT
	610104H7_B	CAAGAGGCCATCACTTTAGC
	610112E3_B	GTCTCCCCGCATAGTTTTTC
30	610104F11_B	AGCCTGGCCTCAGCAGCTTT
	700032C7_B	TCTGTAAGCACGTAGCGGTT
	880032H10_B	CGAGGTGACTTTAACGGAGA
	880022D1_B	GCATCCAATGCCTCTAGCAA
	960050F5_B	GCTCCCCCTGATGAAACTTT
35	610091D3_B	CATAACCGTTGCATGTTGGC
	700011E2_B	CTGGAGGGGAGAATTCTAAG
	610114A7_B	CCCGTTTCTTAAGGTTATCG
	700034B3_B	CACAAAACCATCCTCTTCAC
	960050D7_B	GGATCTGCATCGAGAGAAGT

610093C7_B	GCAGCAGCATACCAATTTCC
610232H7_B	CCCTCCCTCTTCCTTAGTTA
880012D5_B	AGAATGGTTTTCGGCCATCC
610151B8_B	CTACCGTGGAGATTCGTGAT
700054G5_B	GCTTGTCATTGGTAATGCGG
880032A6_B	GATGCGATTACACGATCCCA
700033G7_B	CCGCCATCCTTTTGGATAAC
700052F4_B	GAGTTCCTGCTCAGGAAATC
700053E3_B	CCACGTATTTATGGGACCGT
610202G8_B	CAACGGAACAAGCTATGTCC
700052F2_B	CGCAAAGCGAAGAGAGCTTT
610083A6_B	CCACCAGACTGCTTAACGTA
610131D6_B	CTAGACAAACGAGCTCTTGG
700032E10_B	GTATTCGGCACGAAGCCAAG
610081C9_B	CCCCATATACACGGACACTT
960050A6_B	CTTAGGGAAAGATCCCCAAG
880032H4_B	ATGCGTGTTCGTATGGCTCT
610281H1_B	GCTGCTCGAAACCTACAGAC
610181F3_B	AGGTGCTTCTTTCTCTCTCC
960050C1_B	CTTGCTTGCGGGCAACCATT
610242B10_B	GATGATAAGAAGACTAGCGG
610241D11_B	CATCACGATTAGAGGCTCCA
610103E9_B	GGAGAGACGTCATGCTTTGT
610252E8_B	GGGTCACCCGTATTTAGGTA
610214H5_B	GGTCCCAATTTCCGTTAGGA
610312F6_B	CGCGGTTTGTTGATGAGCAT
880024F11_B	GCAGCACACAGTAGATATCC
610164E4_B	CCATCCTTCAGAGCTCTTCA
700024E2_B	CTCCAGATGTTAATGGAGGC
610182A6_B	CGGCTCTGGCGCTAGTGCAA
610121E1_B	TGCGAATGGCATGAGATCAC
880033D9_B	CCAGGCTCTAACTTCTCATC
T7	TAATACGACTCACTATAGGG
610164D2_B	CCTGGGCATAGATGAGTGAT
610193H7_B	CTAGGCCCCTTTTTGAATGG
610064C6_B	GCTAATCAAACAGGCAGATA
610202F6_B	CGGATGCAAAGCCATCTCTT
610151D12_B	CTCCGGTAAGAACGCGTTTT
610114G8_B	CGCGATGTTATTAAGCACCC

700033H6_B	GCTTTACAGGGAGTTTGTCC
610202B9_B	GATGGGCATGCAGTGCTCGT
610124A7_B	CCGCTGTGGATTTTTAGGAC
880032D12_B	ACCGGAAATGCCTGTACCGC
610233C10_B	AAGGAAGAGCTTGTGGCGTT
700052D1_B	GCCTTATCTCCAGCAGCAAA
880032G8_B	GAGGCAGAACGAGAATGTTC
T7	TAATACGACTCACTATAGGG
610031B10_A	CGAGCATGCTAAAGGAGCTA
610032D3_A	GCAGCAGAACTCCCTCCATC
610032G2_A	CACAGAGGGACTAAATAGGG
610034B2_A	GTTGCACATCAAGCTGAGCA
610034D10_A	GCTCAAATGCGAGAGGGAAA
610041A3_A	CGGAGCTAGATCAATAACGC
610051G1_A	GAACGAACATCTCCTCTAGC
610052F6_A	CCAGCAAGCTTTCGACGAAA
610053E12_A	GCGATCTCACTTTCCCACAA
610054D6_A	CACTGCGGATTTTGACAAGG
610063H6_A	TCCGAAGAGCGGAAGAAGCG
610071E4_A	GCGTGAAAAAGAACTTCCCC
610071G2_A	CCAGATCGTTTCACAGCAGA
610081D11_A	GGGACTAGTTCGACTCCAAT
610081D6_A	GGGAAAGAGCTTGTTGCGAA
610084D5_A	CAGCACCGCTATAGAAGCAA
610091F4_A	CTCCCTTAGGTCTTTGTCCA
610091G11_A	GACCTTTGGATGCTGCCATA
610093D12_A	GCGCAGGATTTGCTTATGTC
610093G12_A	ACTTCTTCGCCACTGATCGC
610103D11_A	CTACTTGGCGATAGGCCTAA
610104G12_A	GGCCCGAGTTCAAGGAAATA
610111A1_A	GCTCTTTTCGGAGTAGGAGC
610111F1_A	GAACGATTACCAAGATCGGC
610112H2_A	GTTTGGAAGGTGAGGCTCTA
610124H5_A	CCCAAGTTTCTACTAGGAGC
610141F9_A	CGAAAGATCCGTGTGTGGAA
610142E3_A	CGACATCTTTACGTTCGCCA
610143B7_A	CGGCAGGTTATGGCTCTTTA
610143E10_A	CATCAAAAGCATCACCAAGG
610151D9_A	CGGAACGTGAATCTGTTCAG

	610152B5_A	AAACACCCTCTCCTACGAAG
	610153A1_A	TGGGATGGATACTGGGGATA
	610164F9_A	CGATCTCGTCCCTAGAACTT
	610172D10_A	CCGACCACATCATGATTGTG
5	610173A7_A	CCATACGCAGCAAGAGTTTC
	610174G8_A	CTGCAGGGATTGATGAATCG
	610174H3_A	CTGATCATCTTCATGGCGTC
	610182F2_A	GTTGAGTAAGAGGAGAGTCT
	610193F1_A	TGTGGACTGCTCTGTTTACC
10	610201G6_A	CTGATCGAAGCGGATGATGA
	610202D8_A	GATGCTCAGGATGCTTTGAG
	610203A8_A	GGAATGGGTGAATTGCACCT
	610204C7_A	CCAGAAACGCAATACCCTAA
	610233G6_A	CCTGCGCCCACGAAGGATAG
15	610234G6_A	CCTGGAGTATACAGATGGAG
	610274B3_A	GCGCGAGAGAGTATTACACG
	610281F5_A	GAGGCTCTTCGTTATTCGTG
	610283H1_A	TCTTCTTGATCGAGCGCAGG
	610304B5_A	GCAAGCGATGAGAAGTTCTG
20	610312C11_A	GGGAAACCTAGCAGACTTGT
	700011A2_A	CGCTGAGTAGACAAAGCTTC
	700011B8_A	GGAGAGCGAGTGACCTTATT
	700032D12_A	GCAGCAGTCTTCATGACGTT
	700033D6_A	GAAGCGTTAATTGCATGCGG
25	700033F9_A	CTTTCCTCTCCAACCCTCAT
	700034C6_A	CCGAATGAAACTGGTTGCCG
	700034F7_A	CTCCGAGAAATTCCAAGCGT
	700042F9_A	GGATCAGTCCTTCGATACCA
	700044G6_A	GGCCTCCCATACAAAAGCGA
30	700052A9_A	GCAGCCATATCACGAATTCC
	700052G10_A	GGGATGTTGCGAGAAGACAT
	700054A5_A	GAGCATGCGGGTATAGTAGA
	T7	TAATACGACTCACTATAGGG
	880012G1_A	CGGCAGCATTGTTACTGTCT
35	880022B4_A	GGAGCTTTGAACAAGATCCC
	880022F6_A	CTGAACGAAGGACTAGTTGC
	880022G11_A	GATGCTTCTCATCCACTTGC
	880023A3_A	CGTCTGCTTGAAGCGCAAAA
	880023C11_A	CTCCCCCATTATCACTCTAC

880023D1_A	AACTCCTGACACTCCCCTGC
880023F3_A	AGCCACGCTTCAGCATAGTT
880023G9_A	GTCAGGATATGCTGTTGGGT
880032D12_A	GGAGATCTAAAGGACAGGAG
880033A7_A	ACACCCAAACCACCCGTTGC
880033C10_A	GCATTCGTGGTATCCAAGAG
880033D8_A	CAGATTGCCGGTTCATTAGC

To avoid the step which consists in ordering and then preparing the clones by conventional microbiological means, outer and inner primers oriented towards the regions not yet sequenced are defined by the software. The primers thus determined make it possible to prepare, by PCR, a template covering the nonsequenced region. It is the so-called outer primers (the ones most distant from the region to be sequenced) which are used to prepare this template. The template is then purified and a sequence is obtained on each of the two strands during 2 sequencing reactions which each use one of the 2 inner primers. In order to facilitate the use of this approach, the two outer primers and the two inner primers are prepared and then stored on the same position of 4 different 96-well plates. The two plates containing the outer primers are used to perform the PCRs which will serve to prepare the templates. These templates will be purified on purification columns preserving the topography of the plates. Each of the sequences will be obtained using primers situated on one and then on the other of the plates containing the inner primers. This distribution allows a very extensive automation of the process and results in a method which is simple to use for finishing the regions not yet sequenced. Table 3 below gives the names and the sequences of the primers used for finishing *Chlamydia trachomatis*.

Table 3: Directed sequencing by PCR

The list of the primers used to obtain the template and the sequences of the regions not sequenced

during the systematic sequencing. The primers carry the identifier for the connecting clone between the consensus sequences. The name of the primer is followed by "e" for the primers used for the preparation of the template. The primers used to obtain the sequence have "i" as the last letter of their identifier.

T7	TAATACGACTCACTATAGGG
Ct610091E12_Ae	CTCCATTCCGAACTGCAGAA
Ct610281H2_Ae	CGTGCCTGGCTTTCTTTTGA
Ct610051D10_Ae	CGCGGAGGTTGATTGCTAAA
Ct880014G12_Ae	AATCCATCACCTCTGGAGGA
Ct960050C5_Ae	GTTACGACCATATGGAGGAG
Ct700033B5_Ae	CCAAGGGAATGGGTTTTTCG
Ct960050E6_Ae	CTCCTATCGTTTGCTCAGAG
Ct960050C1_Ae	GCAACAGAAAACACTCCGCT
Ct610093G12_Ae	CCCGCAGATAGGGAAAGTAA
Ct610172G1_Ae	GGAGGGCCGTTAAGGAATAA
Ct610151B6_Ae	GGGGATCTTCGTTTTGTTCG
Ct880012G1_Ae	GCGTGTCTACCAATTTCACC
Ct610203A8_Ae	GTCACCTAAATAGCTTGGCG
Ct610091F4_Ae	GAGGGAGGATCTGTACAGAT
Ct700044A5_Ae	GAGGAATCCCTAGTTACACG
Ct880032F8_Ae	CTGCCTTAGGGCTTGATAGT
Ct880032B10_Ae	GAGTAGAGGGATTCGAACCC
Ct700031F2_Ae	CGGACTATTCTAGCCTCTTC
Ct960050H5_Ae	GGGCCGTATTCAGACTTTGT
Ct610124A7_Ae	CAAGGCTCTTCCATGTGTTC
Ct610112H2_Ae	CCTATCTGGCAACGAGAATG
Ct960050D3_Ae	GAGCAAGATCTCAGGAACGA
Ct610114B3_Ae	GTTGCTCGTGCAGGGAAAAT
Ct610164F9_Ae	CAAGGGGTGAGCATCCAAAA
Ct610252D3_Ae	CGCCGATAGCTTGATGAAGC
Ct610104G12_Ae	GGGGGCACTAGGAAGTATAA
Ct610173H10_Ae	CTGCTGGTACCGGTTGCGTC
Ct610212H3_Ae	GCGCCGATCTTGGCAATTAT
Ct610183B1_Ae	TAGTAGCAAGAGGAGCGAGA
Ct700033H11_Ae	GAGCTCTGCTTTGCACAACA
Ct880013A8_Ae	CTGGAATTACAAGAGGGGGT
Ct880023H10_Ae	CAAGAACACACGCGATTTCC

Ct610273D4_Ae TAGTACCTGCTGCCGGAGTG
Ct700034G11_Ae CGTCGATAGTGATGAATGGC
Ct610093G3_Ae GAGTATTCGCTGCAGCGGTT
Ct700022C9_Ae CTGGGGTATCTGTTGTTGGT
Ct880013H9_Ae GATCCGCAAGAACCTTAGCA
Ct610254G11_Ae CTAACTTACGGGCGATGCAA
Ct610201G3_Ae CGCGTAGCCAAAGTGAAAGT
Ct700031D10_Ae CAAGAGGAACTTGAGGCTGT
Ct880023A11_Ae CACCCATGCCTCGGATCCTC
Ct610161G5_Ae TATGTACATGGTCTGGCTGC
Ct610282D11_Ae CTACTTCGGTTCCTCTCTGT
Ct700052B3_Ae CTGCAGACATACTTCCAACC
Ct880033E4_Ae TGGAGCTTTCCCGTTCTCCT
Ct880022F1_Ae CGGTTTCGCAAAGTTCGTGC
Ct880032F11_Ae CCTCGAAGCTCCTTCTGTTT
Ct610081G12_Ae GAAAGAGCCAACCAACGTCT
Ct880014G6_Ae CGATTCCCTGTCAGAGTGAT
Ct610041D3_Ae TGTGTGTGTGTGTGTGTGTG
Ct610281A10_Ae GACTTTGCTCTTCGCTCGAT
Ct610222F9_Ae CCTCTTTCCGAAAGATGCCT
Ct610131B7_Ae CGAGCTTTCTTACGACCGTA
Ct610144G4_Ae TTTCTACGCCTCTATCCCAG
Ct880032G5_Ae AGGCTCGAGGTAAAGAGCCT
Ct700044H8_Ae GTGCGTCCTTCTTTACCGAT
Ct610264F7_Ae GGGAAGCATTTCCGTTTAGG
Ct610111F6_Ae GATTCGAGACAGAGGCTTTG
Ct700044E6_Ae GGGACTCGATTCCCTGAAAT
Ct610124F4_Ae GAGAAGTGATTGCGTTCCCT
Ct960050G5_Ae CAAAACAAGGAGCTGTGACC
Ct610142F1_Ae GCTGCCTCTAACGTATGTTG
Ct610244G5_Ae CCGATGGGGATGAGGATTTT
T7 TAATACGACTCACTATAGGG
Ct700042G1_Ae TATCAGCGCTGTTCCACAAC
Ct880032A4_Ae CAGCCTCTGCAAAAAGACGA
Ct610081B5_Ae CCGTTTGGTGCCAGTGTTGC
Ct880033E1_Ae CAGCCAAATTAGGAACGCAC
Ct700033D5_Ae GGCCAAAAGTTCGGTATTGG
Ct880032B5_Ae GGACAAGAAGATCTGGAGAG
Ct610223A1_Ae GCCTTGCTCCCTTTAGTGTA

Ct610112B6_Ae	TCCGGGGACTTCATTCCGTT
Ct610153D7_Ae	CTTTCGCAAGCATTTGCAGG
Ct610182D6_Ae	CCGCTGGTTCTTCCTTACTT
Ct610084A9_Ae	GCTAGAGCTCAAGCTTTAGA
Ct610141E2_Ae	CCTTGGAACACTAGAATGGC
T7	TAATACGACTCACTATAGGG
Ct610091E12_Be	AGACTCTGAATCCACGCAAG
Ct610281H2_Be	GCTACAACACGTGTTTTCCC
Ct610051D10_Be	GCTCTGGAAGCATTTTTCCC
Ct880014G12_Be	GGTGCTTCTGATCAGGTCAT
Ct960050C5_Be	CAGGGTACAAAACCCCTAGT
Ct700033B5_Be	GAAGAGAACCCGGAGATTTG
Ct960050E6_Be	CGGCAGTGATATAGTGAGGA
Ct960050C1_Be	GACAAGCACGAACGGAAGTT
Ct610093G12_Be	ACCCAGAGCAGCCTTCTTAT
Ct610172G1_Be	CCTCTATCTCTACAGCTTCC
Ct610151B6_Be	CCCTATAAGAGCTGGTCTCA
Ct880012G1_Be	GGCCTTTTAGAAAGGAGGGA
Ct610203A8_Be	GGAAGATCCAACTTTCCGAG
Ct610091F4_Be	CTGGATCAGTCAATTGCTGG
Ct700044A5_Be	CCGGATCCAACTCGAACTTA
Ct880032F8_Be	GGCCTCGGAAGCATCTAAAT
Ct880032B10_Be	CCTCCATGGCGGAGAAATAA
Ct700031F2_Be	GTGGGAACATTCACGTTAGC
Ct960050H5_Be	CCTTCAGGAAGATCTCCTTC
Ct610124A7_Be	CACTCTCTCGGATTTGGGTT
Ct610112H2_Be	CCACTGCATACTCATTCCTA
Ct960050D3_Be	CATCCTGACAATAGCTGACC
Ct610114B3_Be	AGTCACAAGATTCGGTCCCC
Ct610164F9_Be	GCGCCTCTGATGATCAGAAA
Ct610252D3_Be	GAGAGTCTAACATTCCGCTG
Ct610104G12_Be	TCAGGGACACCCCTGACACA
Ct610173H10_Be	GAGAATATCCGAGTTTTGCC
Ct610212H3_Be	AGCAAGATAGTTTGCACTGG
Ct610183B1_Be	GCGAGGACGGGTCTTTGGAT
Ct700033H11_Be	GGCTTGCTTGAATACGCAGT
Ct880013A8_Be	CATCGCTAGCTAGAGTCTTG
Ct880023H10_Be	TATGGACTATCGAGCGCGCC
Ct610273D4_Be	GGAGGAGTTAAACTCAGGAC

Ct700034G11_Be GTCTTCGACTTAGGAGGAGG
Ct610093G3_Be GATTACGTTAGGGTCTGTGC
Ct700022C9_Be GGGTAGCGTCTATGCAAAAG
Ct880013H9_Be GATTACGTTAGGGTCTGTGC
Ct610254G11_Be GCATGCTTCTCTGGTTGTTG
Ct610201G3_Be CGCCACCTTTAAAAGCAACG
Ct700031D10_Be TACGGAAACACTCTCCTGGA
Ct880023A11_Be CAAGAGTGCTGACCTTTCCT
Ct610161G5_Be GCGGTTAGTTTAGTGGTACA
Ct610282D11_Be GCTCAGGAAGACTTAGCGCA
Ct700052B3_Be CCATCCCCAGTACAACTATG
Ct880033E4_Be CCGTGTATGGACGATGATGA
Ct880022F1_Be GCCATTCCCAGCTTAATGGT
Ct880032F11_Be AGAGCTGGATTGGGATGTTG
Ct610081G12_Be AGAGCTGGATTGGGATGTTG
Ct880014G6_Be CCCTTATTAGTCGTTGCAGC
Ct610041D3_Be GTCTACAGTCTTAGTGAGGG
Ct610281A10_Be CCACCTGATCCAGATGATGA
Ct610222F9_Be CTCTTTGCTCTTGGAGTTGC
Ct610131B7_Be CAGGCAGAAGAAAGCAGGCT
Ct610144G4_Be GTCTCGGAAGAATATGGAGC
Ct880032G5_Be CCCTATTTAACCCCTCCTCT
Ct700044H8_Be GAGGAATCCCTAGTTACACG
Ct610264F7_Be GATCTGGCAAGCGTAGGAAA
Ct610111F6_Be TCCATCGGATTGCTTATTCC
Ct700044E6_Be CTTCCTCACGTCACTATCCT
Ct610124F4_Be CTTCCTCACGTCACTATCCT
Ct960050G5_Be CTGGTGAGTAGGGTCCATAA
Ct610142F1_Be CCCTACTCTACGCCGATTTT
Ct610244G5_Be GCAACCCACATCTTTCCAAC
T7 TAATACGACTCACTATAGGG
Ct700042G1_Be CCTATCTGGCAACGAGAATG
Ct880032A4_Be TTGTTCACTGTGGGCCGTTT
Ct610081B5_Be TCGTCACTTGGGGAAACTCA
Ct880033E1_Be CCTGTTGTATTGGTCTTCAG
Ct700033D5_Be GTCTCCCGAAGATCTCATTA
Ct880032B5_Be GCAAGGCTTTCGACAAACTC
Ct610223A1_Be GGTCCCAATGTATCACGTTT
Ct610112B6_Be GTCTCTGGATGCAGTTTCAC

	Ct610153D7_Be	CATCGGACTGTAAATCCGAC
	Ct610182D6_Be	GTCCTCGTGGATACCTTAGA
	Ct610084A9_Be	CAGAGCCAGCTTTAAAGAGC
	Ct610141E2_Be	CAGGCCCAAATACCTACACA
5	T7	TAATACGACTCACTATAGGG
	Ct610091E12_Ai	GGGTGTATCATTGATCTCGC
	Ct610281H2_Ai	CTACTTGGCGATAGGCCTAA
	Ct610051D10_Ai	CTCCCGACTTCTCTCTAACA
	Ct880014G12_Ai	GGTAAGGCTGCTTGTGTGTA
10	Ct960050C5_Ai	GTCCCGAAGATTGAAAGAGG
	Ct700033B5_Ai	CCGATTCTTTCCAAACGACG
	Ct960050E6_Ai	CTTGAGTAGAAACGTCCTCT
	Ct960050C1_Ai	GGGAAACCTAGCAGACTTGT
	Ct610093G12_Ai	ACTTCTTCGCCACTGATCGC
15	Ct610172G1_Ai	TACAAGAGGTCTTGCGACAC
	Ct610151B6_Ai	GCAGCAGCATACCAATTTCC
	Ct880012G1_Ai	CCTGGGCATAGATGAGTGAT
	Ct610203A8_Ai	GGTCATAGCTGTTTCCTGTG
	Ct610091F4_Ai	CTCCCTTAGGTCTTTGTCCA
20	Ct700044A5_Ai	GCTTATCCACAATCATGGGG
	Ct880032F8_Ai	GATGGGCATGCAGTGCTCGT
	Ct880032B10_Ai	CCGACCTATTGCTCCCAAAG
	Ct700031F2_Ai	GGAGGAGTTCTGAAACAGCA
	Ct960050H5_Ai	CGAGAGAGAGAAGCTGTCAT
25	Ct610124A7_Ai	GTCAGGATATGCTGTTGGGT
	Ct610112H2_Ai	GTTTGGAAGGTGAGGCTCTA
	Ct960050D3_Ai	CGATCCTACTTCCGTAGATG
	Ct610114B3_Ai	CCGAAAGGATCGCTCGATTA
	Ct610164F9_Ai	GATATCGCTCCTATGCTGAC
30	Ct610252D3_Ai	CCAACCCGAAGAATCTCAAG
	Ct610104G12_Ai	GGCCCGAGTTCAAGGAAATA
	Ct610173H10_Ai	GAGGAAGCGCTTTTCTTTGC
	Ct610212H3_Ai	GGAGTGGTTCTTGTCAGTAG
	Ct610183B1_Ai	GGATACTAGCAGGTTCGTGT
35	Ct700033H11_Ai	GTTGGATGCTGTGCCTATTC
	Ct880013A8_Ai	GCGTAGATTTACTTCCGGAG
	Ct880023H10_Ai	AATCGGCGACGTTCTCCGCC
	Ct610273D4_Ai	CTGCCTCCTGCGGCTCTTCT
	Ct700034G11_Ai	CAGTAGACGATACACCAGAC

Ct610093G3_Ai	TCTCGCAAACGCGCTAGAAT
Ct700022C9_Ai	CCAGACTCCTTTGCAGAGTT
Ct880013H9_Ai	CAAGACCCCTATCCTGACAA
Ct610254G11_Ai	CCACTCTTTCCTATCCCAGA
Ct610201G3_Ai	GGAAACTCAGACAGTAACGA
Ct700031D10_Ai	GTAAGCCTTGCGCATTCTTC
Ct880023A11_Ai	CGCCTAATCCTCGACTACAT
Ct610161G5_Ai	GGAGATCACCTAACCTAGGA
Ct610282D11_Ai	TGATCTGGCCTGCTATCGTG
Ct700052B3_Ai	CTAAATCTGACATGCCGTGC
Ct880033E4_Ai	CCTCCAAACAACCATCCTGA
Ct880022F1_Ai	CGTTCCATGCTTCAGCTGTT
Ct880032F11_Ai	CTTCTAAGAAGTTGGTCGGG
Ct610081G12_Ai	GGTGTAATCCCATTAGTGAC
Ct880014G6_Ai	GCGCACCACCCCTATTTTTT
Ct610041D3_Ai	TGTGTGTGTGTGTGTAGACC
Ct610281A10_Ai	ACTTCGCCTGATGGAGAGCG
Ct610222F9_Ai	CTTTGCTCTGCTATGGGGAT
Ct610131B7_Ai	CAAGAAGCCTTGAGCCTTCA
Ct610144G4_Ai	CAAGCAGCCATGAGACTATC
Ct880032G5_Ai	CGAGTAGTGGTTCAAACGAC
Ct700044H8_Ai	CCGATATCTCCCTTAGCAAC
Ct610264F7_Ai	CGTAGTGAAGCAGCGAGTCT
Ct610111F6_Ai	CTCCCGAGGTGATTTAAAGG
Ct700044E6_Ai	TGTATGAAGTTGCTTGGGGC
Ct610124F4_Ai	GATGTCGGCACCTTCAGAAA
Ct960050G5_Ai	GCCATACATGCGATGAGCAA
Ct610142F1_Ai	CTACAGCTGGAGGGATAGTT
Ct610244G5_Ai	GAGATGGGAGTTTCTGAACC
T7	TAATACGACTCACTATAGGG
Ct700042G1_Ai	GTAGGGGATCAGCGGTTAAA
Ct880032A4_Ai	AGTCCCGCGGCGGACTTTCT
Ct610081B5_Ai	TTTTACTGCCTCTGTATCGC
Ct880033E1_Ai	TCAGGCCTTCCTACTTGAGC
Ct700033D5_Ai	CGCAAGAATCACAAGTGTCG
Ct880032B5_Ai	GTGAATGAGCAGACAAAGCG
Ct610223A1_Ai	CTAGCATTGTAGCTGCCAGA
Ct610112B6_Ai	GGACAGCAGCTCGAAGAAAT
Ct610153D7_Ai	CCTGTAGTCCCTGAGAGATT

	Ct610182D6_Ai	GCCTCTGGAATCTTCCTTGT
	Ct610084A9_Ai	CGAGAACCACTTGTTTACTC
	Ct610141E2_Ai	CTTCCCTTCATAGATCAGCG
	T7	TAATACGACTCACTATAGGG
5	Ct610091E12_Bi	CTGCCGCTGGAGCTTGTGAA
	Ct610281H2_Bi	GTCTCCCCGCATAGTTTTTC
	Ct610051D10_Bi	GAAACTTTGCGCAGGATTCC
	Ct880014G12_Bi	GCTTTGTGGAAACATCGAGC
	Ct960050C5_Bi	GGACCCTTATCTAACGTGAC
10	Ct700033B5_Bi	GACTCCAGAGTTACAGCAAG
	Ct960050E6_Bi	CTGTGCCCTTTATTACGTCT
	Ct960050C1_Bi	CTTGCTTGCGGGCAACCATT
	Ct610093G12_Bi	CAAGAGGCCATCACTTTAGC
	Ct610172G1_Bi	AGCCCAAATAGCCAGAAAGC
15	Ct610151B6_Bi	AAACACCCTCTCCTACGAAG
	Ct880012G1_Bi	CGGCAGCATTGTTACTGTCT
	Ct610203A8_Bi	GGAATGGGTGAATTGCACCT
	Ct610091F4_Bi	GCGCGGAAGATTTCCAAACT
	Ct700044A5_Bi	GAGCAGCATACCGTAGGGCG
20	Ct880032F8_Bi	AGCCACGCTTCAGCATAGTT
	Ct880032B10_Bi	ATACCCGATCCTTCCAGCAG
	Ct700031F2_Bi	CCATGGTGAAAGTCTTTCCG
	Ct960050H5_Bi	GCACTTCGTTTTCCAGTGCT
	Ct610124A7_Bi	CCGCTGTGGATTTTTAGGAC
25	Ct610112H2_Bi	GCATCCAATGCCTCTAGCAA
	Ct960050D3_Bi	CGGGATTTGCAAACTGCACA
	Ct610114B3_Bi	CCTGCTCCTGCATTAATGGA
	Ct610164F9_Bi	CGGCGGTTTGACGATTTTCT
	Ct610252D3_Bi	GTGCTGCCAATCATTTTGGC
30	Ct610104G12_Bi	AGCCTGGCCTCAGCAGCTTT
	Ct610173H10_Bi	ACGAAACGATCCGAAACGAC
	Ct610212H3_Bi	GAGGAGGGAGAGCATTATCA
	Ct610183B1_Bi	TCGTTCCGCATGCTCCGTTG
	Ct700033H11_Bi	GGTAGCTGCGTTGGTAAATG
35	Ct880013A8_Bi	GGCAATCCGACCTCATCTAA
	Ct880023H10_Bi	TCGGAATCGGATACCTAGCG
	Ct610273D4_Bi	CTGTAGCTTTGGAAGCTGGA
	Ct700034G11_Bi	GGAAATCGGTGACGGAAGTT
	Ct610093G3_Bi	CACTTCTGTGGGACTCCCTA

	Ct700022C9_Bi	GCAACCTTCTCCTCCAATGG
	Ct880013H9_Bi	CACTTCTGTGGGACTCCCTA
	Ct610254G11_Bi	TCATCTCTGCGGAGCGAGAG
	Ct610201G3_Bi	GAGTTCCTTCAGAAGAGGCA
5	Ct700031D10_Bi	GCGACAATTCTAACAGGAGG
	Ct880023A11_Bi	GGGGGGCTAGCTATTCTTTT
	Ct610161G5_Bi	CTGTAGGGAGGTTATAGAGC
	Ct610282D11_Bi	CTATTCATGACTGCGTTGCC
	Ct700052B3_Bi	CGCCTTCCTTCGTGGATTAT
10	Ct880033E4_Bi	CAAGCTTCGGTTCACAAAGC
	Ct880022F1_Bi	GGACGTACTGCGTCTCTAAA
	Ct880032F11_Bi	CCTGGGATGCAACCACAAAA
	Ct610081G12_Bi	CCTGGGATGCAACCACAAAA
	Ct880014G6_Bi	GGCAGTACGGCAAGTAATGA
15	Ct610041D3_Bi	AAGCAGGATAACGGCGACAA
	Ct610281A10_Bi	GAGCAGGAACGGGAGGAGAT
	Ct610222F9_Bi	GGGCAGGAGAAATCCCATCT
	Ct610131B7_Bi	GTTTCTAGCGTTCGTCTTCG
	Ct610144G4_Bi	GTGCTCACTATGCTTTGTGG
20	Ct880032G5_Bi	GCTACCAAAAATCGGTGGTG
	Ct700044H8_Bi	GCTTATCCACAATCATGGGG
	Ct610264F7_Bi	AGGAGGAGTTAACGCTGGAT
	Ct610111F6_Bi	GCTTTTGGTCCTTTCTCTGT
	Ct700044E6_Bi	CCTGCAGAGCAATACAGTGT
25	Ct610124F4_Bi	CCTGCAGAGCAATACAGTGT
	Ct960050G5_Bi	CGTGCAAAACAGGATCGTGA
	Ct610142F1_Bi	CCATCCTAATGCTCCCCATA
	Ct610244G5_Bi	GCAAGCTTTTTTCCTCGACG
	T7	TAATACGACTCACTATAGGG
30	Ct700042G1_Bi	GTTTGGAAGGTGAGGCTCTA
	Ct880032A4_Bi	GGATAAAGAGACCTCAGGCT
	Ct610081B5_Bi	GCGCTCAAAGTACAAGTCGG
	Ct880033E1_Bi	CGGATCTCGCATCAGCAAAA
	Ct700033D5_Bi	AGCACTTCCTCAAGAAGTGC
35	Ct880032B5_Bi	CGAGAGTCATAGGACGTAAG
	Ct610223A1_Bi	CGGTTTTCTCTCTCGTCAGC
	Ct610112B6_Bi	GACTGTCTGCAAAAGGCTCT
	Ct610153D7_Bi	CCTTACGGATACGTTGGTTC
	Ct610182D6_Bi	CTATGGGAGATATCCCAGGC

Ct610084A9_Bi GAAGTCCCTCCTACTAAGGA
Ct610141E2_Bi CGCTCTGGCAGTTTTTACTC

Finally, a number of contigs exist in a configuration where one of their ends is not linked to any other contig end (Figure 3b) by a connecting clone relationship (a connecting clone is defined as a clone having one sequence end on a contig and the other end of its sequence on another contig; furthermore, this clone must be derived from a plate or a subset of plates with adequate naming quality). For the *Chlamydia trachomatis* project, this particular case occurred 37 times. Two adjacent PCR primers orienting the synthesis of the DNA towards the end of the consensus sequence are defined for each of the orphan ends of the consensus sequence. The primer which is closest to the end of the sequence is called the inner primer whereas the primer which is more distant from the end of the sequence is called the outer primer. The outer primers will be used to explore the mutual relationship between the orphan ends of the different contigs. The presence of a single PCR product and the possibility of amplifying this product unambiguously using the inner primers evokes the probable relationship between the contigs on which the primers which allowed the amplification are situated. This relationship will be confirmed by sequencing and will allow the connection between the orphan ends of the consensus sequences. This strategy has made it possible to obtain a complete map of the *Chlamydia trachomatis* chromosome and then to finish the project.

Quality control

All the bases not determined with certainty in the chromosomal sequence were noted and the density of uncertainties was measured on the entire chromosome. The regions with a high density of uncertainties were noted and the PCR primers spanning these regions were drawn and are represented in Table 4 below. The name of the primer is followed by "E" for the primers used for

the preparation of the template. The primers used for the sequence have "I" as the last letter of their identifier.

Table 4: Primers used for the quality control

A1T7	TAATACGACTCACTATAGGG
B195359E	CGTCAGAATGCTGATGAGGA
C1974679E	CCGAAGAACGAGCGATCTAT
D1250872E	CCTTCCCAGTAAACGGACAA
E1912887E	CCTTTCTTAGGGCGCCAAAA
F1308701E	CGCTTTTTCCCTACATGCTC
G177537E	CCACATTCGATAGCAGCTTC
H1855471E	CCACATCACGTTTCAGGTCT
A2639394E	AAAGCTCCACCAACAGCTGC
B2344510E	GAGAGAAATGCTTCCTCAGC
C2751717E	GATTCTTTCGTAGCGAGGAG
D2394881E	CGAAGCAAGATCCACTGCAT
E2876573E	CCAGGTGGTGAAATTGGTAG
F243235E	GGGGTAGGGCATAACGTTTC
G2425806E	CGCTTGGCTGTTGTGTTCAA
H2533344E	GATCTCCGATCGCTTTACGA
A3332578E	CAACTGTCTCTCCAGTGAAC
B3159918E	GTCGCCCCAAAAGCTTTTGT
C3373238E	GGTGCGGTCTATGCTAAAAC
D3636558E	GATACCGCATACGCTAAGTC
E3804122E	GGCATTTGTGGCATGGTATC
F3341899E	CTGCCATCGGGTAGAAATTG
G3992660E	GGGGGCGTTACACTGAATTA
H3829404E	GGGGGACTCCTATACCTATT
A4825952E	CTCGGAGTTTGGATTTAGGG
B4898826E	GGAGCGATTCAGTTAGCACG
C4934481E	GCTGCTGTCTCCGACGAGCA
D4881367E	CAATTGACTGAGCTGGGCTT
E4928559E	GGAGGAGTTACTCCAGGAAA
F4843347E	CACACCAATGCGAGAACGTA
G4522271E	GGAGCTGGAGAGGTTTATTC
H4743899E	GCTTGGCGGATATCTTCTTG
A5792976E	GAAGAGAGCTGTTGTGAACC

B5551397E GCGCAGAGCTTGGAATATCA
C5712030E GCTCCGGCATTTATTAAGCC
D5381355E GGCAAGGAATACCTTGCCTA
E5891241E CTCCTCGAACTGCAGTTTTC
F5965234E GCTGCAGGGATTACTGCTAT
G5878715E CTTAGCAAAGCAGCTACACC
H5971424E GTCTACAAGTTAGGGAGCGT
A6407923E CCCTACAGCAGGGAAAGAAA
B6497449E CGACGAATACACTCTTCTGG
C6899955E GACCAAACGTAGGACAAGAC
D6733207E CAGAATGGTTCTTCCGTGTG
E658413E CGAGGCGATGAGATTGATGA
F6580376E CTTGTCGTTGAAGAAGGGCA
G6108525E CTGGAATTACAAGAGGGGGT
H6678885E ACCCAGGCGCTTCAACAGGA
A7259525E CCCGGTGCCTAATTACACTT
B7151905E CAGTGGCAAGCACGTTTAGA
C7617717E CGCTTACACCTTGCGACAAT
D71005024E GGAGTTAGGACGATATTCCG
E7896502E CCACATACTTCTGCCTCGAA
F7172269E GAGGATCCAAACCTTTCTGG
G7248122E GGGAACGCAATCACTTCTCG
H7139900E GCCGAAGCTTCATCAAAAGC
A8603202E CTGCAATCTGTCTCGACCAA
B8127845E GGGGAATAGCAAGAGGTTTC
C8698847E GAATCTGCCAGCGCAACTAT
D8913758E GAGGAACTCTTGTCCTTGGA
E8774399E GAGCTCTGCAGGATTGTAGA
F8875472E GAAGGGAGTCTCTTCCAGAT
G8720171E AGAAAGACGCGGCTTTGGAA
H8704740E GAGCTTTCTCAGTCAGAGAC
A9T7 TAATACGACTCACTATAGGG
B9973181E GGCTAGTGCCTTAGGAATGC
C9635222E CTCCCATTGCCGAGAGATAA
D9903653E GAATGGGCAGCATTCTCTTC
E9198064E GACGACGAAACCCAAAACTC
F949541E CATTTCTGCGCATCGTTCCG
G9173914E TCAACGGAGCAACCAAGGAC
H9941526E GCATCTTCCTCATACAGGTG

A10975457E CAGGATCGTTAACAGGGAGA
B10718154E GTGTTAGCGGAGCCTTCTGC
C10198997E CGTTCCCTGGGAAACTTCTT
D1050528E GGCTCGAACTCCACTCATTA
E10534230E CCTGCTCTCCCCAATTTTTC
F10401347E CTGGGAGATCGGTTGTTACT
G10762081E CCACGACTGTCATAGTTCCA
H10627427E GTCCACAAGCTTTCCAGAGA
A11703157E GGGTAGAGCCGTTAATCGTT
B11506316E CAACACTGCCCCACATTATC
C1193701E GGGCGCAAGATTGAACATCA
D11137080E GCCGCCGATTCTCTCGGAAA
E1144034E GCTCGCGCTACATCATAATC
F11632522E CCGCAAGAAACTTTAGCCAC
G11274105E GCTGAAGCATGGAACGAGAA
H11141900E GCTCCAAAACGTCTCGCTAT
A12124269E GAAGAGAGAAGGTGGTTCCA
B12768061E GGGCTATTATGCTCCGTCAA
C12737099E GGCCTTCAATTCCTTCTTCG
D12931806E ATAGGGGGAGCTTGCATTGT
E12864015E GCAGCGGTATTGCTATCGAA
F12731457E GAGAGGAGAGTAGCAGATTC
G12898269E GCGCCGGACAGAATCAGAAT
H12523503E CCGTCAGGTCTCTTTCTACA
A1T7 TAATACGACTCACTATAGGG
B194589E CTCCTGTAACAGAAGGAGTC
C1973481E CTTCCCGGATAAAGGAGGAT
D1249778E CGCTCCACTTCTTTTCCATC
E1912164E GGCTCTGACTCAACAAAACC
F1307817E GGATCAGGAGCTATGCAAAG
G177137E CTCGCTTGATCGATGAGAGA
H1855071E CAGCTCTGGGAAACCTATCG
A2638994E GACATCACAAGGCCCTCTAA
B2344110E CGGCCAGCTTCGTTTGTAAA
C2751317E CGCTTACGTTATGCGGAGAA
D2393524E GCCCTTTAGGGTTGTCATAG
E2876056E CCACGCGGATCTTCGTATTA
F242472E AGAGGATGCTTTAGCTCAGC
G2425390E GTCGGGATCAAGTATATGGG

H2532788E	CGTCGTTTATTGGCAGCTTC
A3331640E	CCTTGGATCGTGTACAGTCA
B3159264E	CTTACAGAGCAAGACGAAGG
C3372838E	CTGTGAAGAGAACAGAGCCT
D3636125E	GCACGGCCATCAATACAATC
E3803359E	GGCAGCTAGAAATCCTACCA
F3341453E	GGAGCACCCTTGTGTCAGGA
G3992050E	GACCTCTTCGTTCTCCTGTA
H3828759E	GGCATCCCAAAAACGTCGCC
A4825552E	CTCCTACAGGGACAGTGATT
B4898424E	GCCTATGCAATCCTACTTGG
C4934070E	CGCATGAGAATATTCCCACC
D4880816E	CCGAGAAAGCTGTCCTTTGA
E4927676E	CCTGAGGATACTGCAAGACA
F4842947E	CTGGTTTGCTCTCTACCCAT
G4521871E	GCAGCAGAGCCAAAACAAAC
H4743499E	AGCTAATCAAGAGGCAGAGC
A5792557E	CTCGCTATGAACTTGCTTCC
B5550937E	GCGGCTTATTGTAGTAGTGG
C5711565E	CCTTCACCACACCTTCAGAT
D5380802E	GCAGGAACTTCTGCAGAGTT
E5890683E	CCCAACAGCTCTTAGCATCT
F5964352E	AGGAGATGAAGCGGTCAAAG
G5878315E	TCCCAATTGGAGACTCCTAC
H5971024E	CATACCGAGGATCCTAACCA
A6407523E	CCTCTTGTCCTCCAATAGCA
B6497049E	AGTAGCTGGGCACAGAGGGG
C6899555E	GGCAGATGTGCAGATTGCTA
D6732807E	CTCGAGATAGCAGGTCGATA
E658013E	CCACCGCAACAAGAACTTCT
F6579976E	GCTTCTGTAGCTGCATCTGT
G6108125E	GCTCCTCCCATAACAGTAGA
H6678485E	CAACAAGAGCAACGATCCCT
A7258673E	CCGCAAGAAGGGGATGAGGG
B7151505E	GAAGTCGGTAGCCCAACTAA
C7617317E	AGGCACGAGCTACTATAGGA
D71004624E	CCGCCAACACAGATGCAAAA
E7896102E	CAACCTCACAGGTAGGAGAC
F7171869E	CCTCGACGTAAGGCAATTCA

G7247722E GCTTCGGAAGACGTATTCCA
H7139500E CCATCATTCTGGAGAGCGTT
A8602802E GCCAAGGAAAGTGATCGGAA
B8127445E GAGAAACGCTCACTAAGCAC
C8698447E CACGCACATGGTTGGGAAAA
D8913358E CAACTCTTGCTCGCTTAACG
E8773999E GATCTACGCCAAGCTAAGCA
F8875072E GTACTATCCTCTCCAGACCT
G8719771E CCCGATTGTAATCGGCAGTA
H8704340E TCCGAGCGTTGGAAAGAACG
A9T7 TAATACGACTCACTATAGGG
B9972781E CGGCCTGTTAATAGGCATCT
C9634781E GCTTGCTGCAATAGAGGGTA
D9903104E AGCTGGTTTAGAGTCTGCTC
E9197333E GTTAGCGTGTGCAAATCTGC
F949141E CGGAAGGCGTAAATGGTTTG
G9173514E CGGAGTTTCTGGATTTGTGG
H9941126E CAGTTTAGATCCAGCAGCCA
A10975057E CATTCCAATCAGAGCCAGAG
B10717754E GCCGCCTTCTATCCTAAAAG
C10198597E CCATAGGTGAGTAAGAGCCA
D1050128E GCCCCATATCCATCATCACA
E10533830E GAGACGGCAATAACGCAGGC
F10400947E GCTAAGCGATCCACATTAGG
G10761681E GGGGCTTTTTCAAATCCGTC
H10627027E AACAATGGCTCCCATAGGAG
A11702757E GGATGCGCTGACTCATATAC
B11505916E GGGATTGCTTGTAACGTGAC
C1193301E GGAGATTGCATGCTGATACC
D11136623E GACCTCGGCCTTTAAAGACG
E1143576E CATTCAACTGCAAGCTGCTG
F11632057E GATAGAGGGGTTCTATGACG
G11273623E GAACCCGAGGAAGAGTCTCA
H11141413E GCCAAGCCATATAGCGTTCT
A12123769E CCTGGTCATAGCCAACTGTA
B12767559E GGCCCCATATTCCAAACAGG
C12736519E CTGGTGCGATCCTAAAAGGT
D12931214E GGCGTGAAGAGAGACACTAT
E12863327E CCGGCGATTTACTTTTTCGG

F12730709E GGGAGCGTCATTTTGGAGTT
G12897869E CCCTAAAACAGCTTCCCTTC
H12523103E GCAGAACTGCCAGAGAAGAT
A1T7 TAATACGACTCACTATAGGG
B195359I GGGAGAGATTCCCGCATTTA
C1974679I ATCTAGGAGAGAGCTTGTGC
D1250872I CCTCTGGCTTACCAGAAGAA
E1912887I GGAAGGAAGGCCCCGAGTAT
F1308701I GAGTACGCTCTTGTTGCACT
G177537I CTTCATCTTGGGATTCCAGC
H1855471I GGATGTAGGCTCCTAGAAAG
A2639394I CACACGCTCCAACAATAGCC
B2344510I GCTCACAATAGCCAAGAACG
C2751717I GGAACAGAGCCAGATAGTCA
D2394881I ACCTTGGAGATCACACCCCT
E2876573I CTGGATTTAGGATCCAGTGC
F243235I CCCTTTGTCCCTGCAGATAA
G2425806I CGTAGCAACAATCTCCCCAT
H2533344I CCCTATTTAACCCCTCCTCT
A3332578I CAGAGCTCCCTACAACATAC
B3159918I GTTAAAAGATCCCGCTCCCA
C3373238I GACTCTAGACGAACTGTCAC
D3636558I CCCTGCTGTTAGGGTATCAA
E3804122I GGGTTCTCGTCCACGCTGAT
F3341899I TACATCTGTTGTGGGTAGGC
G3992660I GGGAAGGGCGTTATATCAAG
H3829404I CGATGGAGAGCTTTTGGTCA
A4825952I GGCTTGAGCTTGCTTAGCAA
B4898826I GGGGGCCCTTGATTTGCGTT
C4934481I GATCTATATCCCCGATAGCC
D4881367I CAGAGAGAGCAAACTCACGA
E4928559I GGGTGGGAGCCAATTTTTAG
F4843347I TACGTTCTGTGAGGAGAACG
G4522271I CGGCATCACAGAAGATGTAG
H4743899I GCAAGGACAAGCTCGTCTTT
A5792976I GCTCTAGTGGGTGTATTTCG
B5551397I GACTCCAGACGTAGCATCTT
C5712030I GGGGATACTGTTTCAGAGGA
D5381355I CTAACGCCTATCGAGTTCGT

E5891241I CTAAGCTGTGGTAAAGCGTG
F5965234I TCTGTCGGCTTCCTTCTGGG
G5878715I GCATATGGAGCGAACTCCTA
H5971424I CCTCGTCTCCAGTAACTTTC
A6407923I GGGTGTGACCCTTCGGGATT
B6497449I GGAGTATTTGGGCACTCCTA
C6899955I CGGACGCAAAGAGCGATCTA
D6733207I GTCCTCCACGGATAAAGAGT
E658413I CCAGATAGTGAAGAGGGAGA
F6580376I TGCGGAAGTGATTACGACGA
G6108525I GCGTAGATTTACTTCCGGAG
H6678885I GCTGCAAAACCACTTCCAGA
A7259525I AAACGCCATGCCCTCCACCC
B7151905I CTTGTCTCAGGACTTCCTTC
C7617717I AGATGCCCCACGTGCAACCC
D71005024I GTAGCTAGGAAAGCTGTGAC
E7896502I CTTGGGTTAAGACAGGGGTT
F7172269I CCTCCGCTAATAACGCTTCT
G7248122I CGCTCCTGTTAGAACCATCA
H7139900I CCGAAGCTTGCCTCTACGCA
A8603202I CTGCAAGTCTCTGCTTATCC
B8127845I GCCACAACTCTCCTTCTTCT
C8698847I AAGTAGAACTGCGCGGTGTA
D8913758I GGGTGTGGAGATAGGAAAAG
E8774399I AGCGGTTCATCCACCAGTAG
F8875472I GGGCTGTTGATCGAGCATTT
G8720171I CGGCGTTTCCAGCCTATTTT
H8704740I CCGGTAAGAGCATCCGTTTT
A9T7 TAATACGACTCACTATAGGG
B9973181I GTAGAGAGCAGAGATCACTG
C9635222I CTAGCCCCCCTCCAAATAAA
D9903653I CTCCAGATCAACCGCGTAAT
E9198064I CAACACGTTCCTGCTTTTGG
F949541I GAAAGCCGATTCTGATCGGG
G9173914I GACTGCACCATGCTCTTCAA
H9941526I GGGGAGCCATTTGTACTCAA
A10975457I CCTTCTCTCTAGATAGGGTG
B10718154I GGAAGAGCAAAAGGCTCGAT
C10198997I CGCGCATACGCAGAGATATT

D1050528I TCCGGGCCAGGAGAAAAACA
E10534230I AGCAACAACCGAGGAATCCA
F10401347I CTCCTGTGTGGAAGGAGAAT
G10762081I GTGAGTACATAGCGACTCTC
H10627427I AAGCCTTCCTTTGCTTGGGA
A11703157I GAAGTGATAACCTGCGCTCT
B11506316I GGGCTGCGATCTGTCTAATA
C1193701I GCTCTCTGCCAATCAAGTTG
D11137080I GAAGCTCCCGACTCTAAGAT
E1144034I GCTGCTTGCAAAAGCCGTAA
F11632522I CGAGCTCCCCAATCATTTGA
G11274105I CACGAACTTTGCGAAACCGT
H11141900I CTCGCCATGTTCAGTCCTTA
A12124269I CGTGAGGACTAGGAAAGACT
B12768061I AGGGAGTATTGGTAGACAGG
C12737099I GAACACGTTTTCCTGGAGGA
D12931806I CGTGTTATTCCCAGTAGCCA
E12864015I CTGTTTGAATGGCTCCTCCG
F12731457I CTGGATCCAGGTTTCTAGAG
G12898269I GAGGAGTAGAACAAAGCTCC
H12523503I CCCTATTGTAGAAGGCTCTG
A1T7 TAATACGACTCACTATAGGG
B194589I GTCCTTCTTGTGAAGAGACC
C1973481I TCCCTAGTCTCTTTGGGGCA
D1249778I CGTTGGGATAGAGGAAACAC
E1912164I CTCTAAGCCCTGACACATTC
F1307817I CTCAAGCAAGCTCTCGTTCT
G177137I GCTGCTCACGTAAATGCACA
H1855071I GCCTTGGCTAACTATTCGGT
A2638994I CAACGGTACTCTTCTGACCT
B2344110I GCTATTGTGGGATGCGTGAA
C2751317I CGTCTTGATCATATCGTCGG
D2393524I CGCCCATGAGATTCAGTTTC
E2876056I GGCATTCTCGCATTTCTTCC
F242472I GAGCGAGAAGAACTTCTCTC
G2425390I GGATCACACGGGCCTATTTT
H2532788I GAGTCTAGAAGCTGCGATTC
A3331640I GTGCAACGAATAGTGCCTTC
B3159264I GGACAAGCTCGAGAAGTCTT

	C3372838I	GCGCTTCCTTCTTCTAGAGA
	D3636125I	GAAGGTTGGCCAGCTTTTAG
	E3803359I	CCCTTCGGAACCTCTTTATC
	F3341453I	GGGCAGCAGTAGCTACAATA
5	G3992050I	TCTACCTGATCTAAACGCCC
	H3828759I	GCGACCTACGAATATACCAG
	A4825552I	ATGGGGCACCAAGCCGCCTC
	B4898424I	GTACTGTCTATGGCTTTGGC
	C4934070I	ACTCTGCCAGAGAGACCATA
10	D4880816I	TCTGCCACAGCAGAACAGCA
	E4927676I	CCTAGCATTGGCGAAGAGAA
	F4842947I	GGAGGCTGTCGTTACTGAAA
	G4521871I	GTTACTGTGGCCACCAGTTT
	H4743499I	TGGGAAGGGTGGTTTAGGAA
15	A5792557I	CCAAGCGGCTTCTAAACACA
	B5550937I	CGTGACCGATTCAATTCCTC
	C5711565I	CAGGGATTACAGGGCACTTA
	D5380802I	GGAGGAGAAAGTCGCTTTAG
	E5890683I	CTCTGACCATTCATCAGGGA
20	F5964352I	CTGCTCCACAAAACAAGGAG
	G5878315I	GACTGGTCTGAGGAAGGAAA
	H5971024I	CCTTGGCTTTTGGATGCGTT
	A6407523I	AACTCGTTGGAAGAGAGGTG
	B6497049I	TTCCCGGTACGTCTTAACAG
25	C6899555I	GGAACTCAAGGAGCACCTTA
	D6732807I	CATCATCCCCAGAAGCCATA
	E658013I	GCTCCACAAAAGTTGATCCC
	F6579976I	CTTCCAAAGGTAGGAGCTGT
	G6108125I	GCATCATCCTCGCAAACCAT
30	H6678485I	CTTCATGACACCCTTGGGAA
	A7258673I	GCTCGTAGCAGATCTTTGTG
	B7151505I	ATGCAAGCTCTGGTACACTC
	C7617317I	GGAGTGATATTAACGGCTCG
	D71004624I	CCAGCTGTAGGGATATATGC
35	E7896102I	ACACAATACGTCTCCTCGCT
	F7171869I	GTTACGAACTATGTGTGCGG
	G7247722I	TCCAGAGGCTGTTGCTAACC
	H7139500I	GGTAGAGGAGTAAGTGTTGG
	A8602802I	GCCTTTTAGAAGAGCGTGTG

B8127445I CTTCCTCATCTCCGGAACTA
C8698447I CGCCATACCCGTGACAAAAA
D8913358I GTCCTGTTCCTGTAGCCGTT
E8773999I GTGCTCTCACCATTCGTTGT
F8875072I GTAGGATGGAGAGCTATTCC
G8719771I CGCTCAGCTTCCTCTTCGGC
H8704340I GGCAAACTCATTGCTGGAGA
A9T7 TAATACGACTCACTATAGGG
B9972781I CCCGAGCTCTTTCTCTTTCT
C9634781I CTAACGTAGCTGAGGAAACC
D9903104I TCTCTCGTCATGACCACCTC
E9197333I AATGGAGCCGATCCAATGCC
F949141I GCTTGTGCGCTGTTTCAATC
G9173514I GGAGCCTCTTGTATTCAGCA
H9941126I CCAACTTAGCTCCAGCATCT
A10975057I CTCGTCCCTAGCTTCTTGAT
B10717754I CGCCTGAAGAAACTCTCCTT
C10198597I GCTGAAAAAGCTCCTTCTCG
D1050128I AAGTCTTCTGCAGCACCACC
E10533830I GACTACAGCGAGCAGAGATA
F10400947I GCACAGATTGTCAGAAAGGG
G10761681I GAGACGACTTACTTGCTTCC
H10627027I CGGGTTGTACAATGATTCCG
A11702757I GAACTCGCATAACCTTTCCC
B11505916I CGCAGCAGCAAAAGCTAAAG
C1193301I CCTCGTCATCGATCACGTTA
D11136623I GGTTCGTTCCCATTTCGAGA
E1143576I GCACTGGCCAATCAGTAGAA
F11632057I GCTGAGCGTTTTTGAGGGTT
G11273623I GCCTGTGACTTCAGTGGATT
H11141413I GCGTCTGGCTGATAGCCGCT
A12123769I GGCCCTTTGTAGGGTCTTTA
B12767559I GTGTCGCAAGACCTCTTGTA
C12736519I CGGCTTAAAATCTGAGGGAG
D12931214I CAGCCCCACAAGTGCCAACA
E12863327I CGTCCTAGATCCTGTCTCTT
F12730709I TGGCATCCAACGATCCTGAA
G12897869I CGCTCCTTCCATCGCATTTT
H12523103I GTAGTAAGCGGACAGCATCT

Data banks

Local reorganizations of major public banks were used. The protein bank used consists of the nonredundant fusion of the Genpept bank (automated translation of GenBank, NCBI; Benson et al., 1996).

The entire BLAST software (public domain, Altschul et al., 1990) for searching for homologies between a sequence and protein or nucleic data banks was used. The significance levels used depend on the length and the complexity of the region tested as well as the size of the reference bank. They were adjusted and adapted to each analysis.

The results of the search for homologies between a sequence according to the invention and protein or nucleic data banks are presented and summarized in Table 1 below.

Table 1: List of the potentially coding chromosome regions and homologies between these regions and the sequence banks.

Legend to Table 1: The probably coding open reading frames are identified with the GenMark software version 2.3A (GenePro), the template used is Chlamydia trachomatis of order 4 on a length of 196 nucleotides with a window of 12 nucleotides and a minimum signal of 0.5. These reading frames are numbered in order of appearance on the chromosome, starting with ORF2 (ORF column). The positions of the beginning and of the end are then given in column 2 (position). When the position of the beginning is greater than the position of the end, this means that the region is encoded by the strand complementary to the sequence which was given in the sequence SEQ ID No. 1.

All the putative products were subjected to a search for homology on GENPEPT (release 103) with the BLASTP software (Altschul et al. 1990). With, as parameters, the default parameters with the exception of the expected value E set at 10-5. Subsequently, only

the identities greater than 30% (I% column) were taken into account. The description of the most homologous sequence is given in the Homology column; the identifier for the latter sequence is given in the ID column and the animal species to which this sequence belongs is given in the Species column. The Homology score is evaluated by the sum of the blast scores for each region of homology and reported in the Score column.

Transmembrane domains:

The DAS software was used as recommended by the authors (Cserzo et al., 1997).

This method uses, to predict the transmembrane domains, templates derived from a sampling of selected proteins. All the regions for which a "Cutoff" greater than 1.5 was found by the program were taken into account.

TABLE 1

ORF	Position	Homology	ID	*Species*	Score	I%
ORF2	0000501-0000208	putative				
ORF3	0003276-0000505	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	379	37
ORF4	0005068-0003242	lipid A disaccharide synthetase (lpxB)	U32786	*Haemophilus influenzae*	285	40
ORF5	0006373-0005126	poly(A) polymerase	AE000123	*Escherichia coli*	552	46
ORF6	0007977-0006619	D-alanine permease (dagA)	U32770	*Haemophilus influenzae*	265	36
ORF7	0008561-0008082	signalpeptidase II	X78084	*Staphylococcus carnosus*	174	36
ORF8	0008995-0008591	YteA	AF008220	*Bacillus subtilis*	157	43
ORF9	0009440-0008979	ORF 168	D28752	*Synechococcus sp.*	318	42
ORF10	0009828-0010430	unknown	Z80108	*Mycobacterium tuberculosis*	324	46
ORF11	0010367-0011254	hypothetical protein (SP:P39587)	U67605	*Methanococcus jannaschii*	152	38
ORF12	0011245-0011916	rRNA methylase	D90913	*Synechocystis sp.*	209	40
ORF13	0012263-0013324	hypothetical	U32691	*Haemophilus influenzae*	367	45
ORF14	0013532-0014413	neutral amino acid transporter B0.	U75284	*Oryctolagus cuniculus*	410	39
ORF15	0014807-0015019	dihydrolipoamide acetyltransferase	L38646	*Saccharopolyspora erythraea*	324	47
ORF16	0014932-0015969	branched chain alpha-keto acid dehydrogenase E2	M97391	*Bacillus subtilis*	577	44
ORF17	0016004-0016501	ORF_o328	U18997	*Escherichia coli*	223	44
ORF18	0016467-0016138	putative				
ORF19	0018190-0017417	putative outer membrane protein	U80956	*Borrelia burgdorferi*	86	36
ORF20	0020521-0018437	ORF-2	D11024	*Shigella flexneri*	642	37
ORF21	0022202-0020814	dnaK like protein (AA 1-660)	X52175	*Chlamydia trachomatis*	2214	99
ORF22	0022602-0022153	ORF, 82 kDa protein	L22180	*Chlamydia trachomatis*	558	89
ORF23	0022795-0022478	heat shock protein	M62819	*Chlamydia trachomatis*	503	99
ORF24	0023183-0022824	GrpE-like protein	L25105	*Chlamydia trachomatis*	580	98

ORF	Position	Homology	ID	*Species*	Score	I%
ORF25	0023394-0023110	GrpE-like protein	L25105	*Chlamydia trachomatis*	373	87
ORF26	0024569-0023394	has homology to putative heat shock proteins of Bacillus subtilis and Clostridium acetobutylicum; ORFA; putative	L25105	*Chlamydia trachomatis*	1999	99
ORF27	0026383-0024641	aminoacyl-tRNA synthetase	L25105	*Chlamydia trachomatis*	3044	99
ORF28	0026640-0027710	ORFB; putative	L25105	*Chlamydia trachomatis*	1298	99
ORF29	0028780-0027725	putative				
ORF30	0029939-0028740	hypothetical protein	D64004	*Synechocystis sp.*	786	46
ORF31	0030721-0030032	putative				
ORF32	0031281-0030520	putative				
ORF33	0031463-0031780	putative	L46591	*Bartonella bacilliformis*	126	45
ORF34	0033356-0031800	putative				
ORF35	0033901-0033314	putative				
ORF36	0034131-0035027	Yer156cp	U18917	*Saccharomyces cerevisiae*	175	32
ORF37	0034988-0035359	F21C3.3	Z71261	*Caenorhabditis elegans*	245	44
ORF38	0035167-0035919	putative				
ORF39	0035923-0036996	putative				
ORF40	0037810-0037013	putative				
ORF41	0038207-0039085	DAPH synthase-chorismate mutase	AF008220	*Bacillus subtilis*	529	48
ORF42	0039196-0039927	arginine binding protein	X67753	*Escherichia coli*	192	44
ORF43	0039923-0040756	putative				
ORF44	0040760-0042007	hypothetical protein MTCY154.05c	Z98209	*Mycobacterium tuberculosis*	663	43
ORF45	0042175-0043116	phophoglucoisomerase-like protein	L40822	*Chlamydia trachomatis*	681	95
ORF46	0042999-0043802	phophoglucoisomerase-like protein	L40822	*Chlamydia trachomatis*	959	91

ORF	Position	Homology	ID	*Species*	Score	I%
ORF47	0044211-0045227	NADP-malate dehydrogenase	L40958	*Flaveria bidentis*	755	42
ORF48	0046072-0045275	putative				
ORF49	0046340-0045975	putative				
ORF50	0046895-0046506	putative				
ORF51	0047955-0046882	membrane protein (arcD)	M33223	*Pseudomonas aeruginosa*	892	47
ORF52	0048585-0048178	putative				
ORF53	0050072-0048630	putative				
ORF54	0050710-0050099	putative				
ORF55	0052439-0050925	dehydroquinate dehydratase/shikimate dehydrogenase	L32794	*Nicotiana tabacum*	142	36
ORF56	0053484-0052348	3-dehydroquinate synthase	D90911	*Synechocystis sp.*	462	39
ORF57	0054536-0053466	chorismate synthase	X67516	*Synechocystis sp.*	801	56
ORF58	0055086-0054595	shikimate kinase II	M13045	*Escherichia coli*	154	38
ORF59	0056350-0055031	5-enolpyruvylshikimate 3-phosphate synthase	U67500	*Methanococcus jannaschii*	355	37
ORF60	0055659-0056084	putative				
ORF61	0056847-0058235	putative				
ORF62	0058444-0059181	dihydrodipicolinate reductase	U47017	*Pseudomonas syringae pv. tabaci*	350	40
ORF63	0059185-0060195	aspartate-semialdehyde dehydrogenase	U67476	*Methanococcus jannaschii*	590	44
ORF64	0060188-0061483	aspartokinase III	U00006	*Escherichia coli*	312	41
ORF65	0061496-0062353	dihydrodipicolinate synthetase (dapA)	AE000609	*Helicobacter pylori*	345	42
ORF66	0062500-0063141	putative				
ORF67	0063429-0063983	hypothetical protein	Y14084	*Bacillus subtilis*	148	42
ORF68	0064628-0064071	putative				
ORF69	0064285-0064656	putative				
ORF70	0064944-0064609	putative				
ORF71	0065395-0067269	unknown	D26185	*Bacillus subtilis*	733	44
ORF72	0067656-0068873	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF73	0068877-0069233	KsgA	Z94752	*Mycobacterium tuberculosis*	156	38
ORF74	0069212-0069721	high level kasgamycin resistance	D26185	*Bacillus subtilis*	306	43
ORF75	0069958-0070455	polypeptide deformylase	Y10305	*Calothrix PCC7601*	272	43
ORF76	0070710-0071006	protein translocation protein, low temperature (secG)	U32727	*Haemophilus influenzae*	90	32
ORF77	0073191-0071086	putative				
ORF78	0074900-0073497	putative				
ORF79	0075463-0074876	homologous to unidentified E. coli protein	M96343	*Bacillus subtilis*	283	34
ORF80	0077088-0075502	o530; This 530 aa ORF is 33 pct identical (14 gaps) to 525 residues of an approx. 640 aa protein YHES_HAEIN SW: P44808	AE000184	*Escherichia coli*	1447	42
ORF81	0077000-0077299	putative				
ORF82	0078089-0077145	integrase-recombinase protein (xerC)	U32750	*Haemophilus influenzae*	495	38
ORF83	0079065-0078154	hypothetical protein	D64001	*Synechocystis sp.*	400	40
ORF84	0081971-0079878	LON protease homolog	U88087	*Arabidopsis thaliana*	1927	48
ORF85	0082639-0083271	putative				
ORF86	0083792-0084850	DnaJ	U58360	*Salmonella typhimurium*	822	42
ORF87	0084876-0086921	putative				
ORF88	0088650-0087313	putative				
ORF89	0087440-0087805	putative				
ORF90	0088400-0088747	putative				
ORF91	0088717-0089265	putative				
ORF92	0089457-0089732	Hpr protein	X12832	*Bacillus subtilis*	128	32
ORF93	0089762-0091447	PTS enzyme I	U12340	*Bacillus stearothermophilus*	671	34
ORF94	0091749-0091435	ORF107	X17014	*Bacillus subtilis*	120	35

ORF	Position	Homology	ID	*Species*	Score	I%
ORF95	0092392-0091745	putative				
ORF96	0093138-0092344	dnaZX-like ORF put. DNA polymerase III	X06803	*Bacillus subtilis*	542	53
ORF97	0094134-0093361	excinuclease ABC subunit A (uvrA)	AE000583	*Helicobacter pylori*	326	36
ORF98	0094637-0094071	excinuclease ABC subunit A (uvrA)	AE000583	*Helicobacter pylori*	487	40
ORF99	0098299-0094628	UvrA	D49911	*Thermus thermophilus*	2090	44
ORF100	0098715-0098113	excinuclease ABC subunit A (uvrA)	AE000583	*Helicobacter pylori*	319	42
ORF101	0100228-0098741	pyruvate kinase	U83196	*Chlamydia trachomatis*	2411	97
ORF102	0101347-0100337	hypothetical protein	D90903	*Synechocystis sp.*	494	37
ORF103	0102210-0101323	YqiE	D84432	*Bacillus subtilis*	471	49
ORF104	0102485-0102210	putative				
ORF105	0104237-0102726	exonuclease VII, large subunit (xseA)	U32723	*Haemophilus influenzae*	634	51
ORF106	0105009-0104254	triose phosphate isomerase	L29475	*Bacillus subtilis*	558	48
ORF107	0105259-0105894	phosphoribosylanthranilate isomerase	U18969	*Arabidopsis thaliana*	300	38
ORF108	0107429-0108460	putative				
ORF109	0108665-0108955	putative				
ORF110	0109459-0109013	putative				
ORF111	0110366-0109704	putative				
ORF112	0111345-0112520	elongation factor Tu	L22216	*Chlamydia trachomatis*	2007	100
ORF113	0112915-0113463	transcription antitermination protein (nusG)	U32754	*Haemophilus influenzae*	313	37
ORF114	0113572-0113994	ribosomal protein L11	D13303	*Bacillus subtilis*	443	59
ORF115	0114020-0114604	ribosomal protein L1	Z11839	*Thermotoga maritima*	528	54
ORF116	0114720-0115253	ribosomal protein L10	Z11839	*Thermotoga maritima*	143	38
ORF117	0115362-0115676	rpl12 (AA 1-128)	X53178	*Synechocystis PCC6803*	254	62
ORF118	0116022-0119795	DNA-directed RNA polymerase beta chain	X64172	*Staphylococcus aureus*	2675	61

ORF	Position	Homology	ID	*Species*	Score	I%
ORF119	0119823-0124010	DNA-directed RNA polymerase beta' chain (rpoC)	U32733	*Haemophilus influenzae*	3486	50
ORF120	0124095-0124988	transaldolase	L19437	*Homo sapiens*	677	50
ORF121	0124873-0125106	transaldolase	U67611	*Homo sapiens*	121	44
ORF122	0126261-0125536	putative				
ORF123	0126328-0126930	putative				
ORF124	0127138-0127785	putative				
ORF125	0127924-0129714	A1 isoform of vacuolar H+-ATPase subunit A	U22077	*Gallus gallus*	1062	45
ORF126	0129720-0131033	membrane ATPase	X79516	*Haloferax volcanii*	790	48
ORF127	0131018-0131629	putative				
ORF128	0131834-0133156	Na+ -ATPase subunit I	D17462	*Enterococcus hirae*	188	34
ORF129	0133075-0133584	v-type Na-ATPase	X76913	*Enterococcus hirae*	110	38
ORF130	0133625-0133999	v-type Na-ATPase	X76913	*Enterococcus hirae*	89	32
ORF131	0133861-0134508	putative				
ORF132	0134638-0137454	valyl-tRNA synthetase	D64006	*Synechocystis sp.*	1763	51
ORF133	0137442-0140276	PknD	Z95209	*Mycobacterium tuberculosis*	452	44
ORF134	0140733-0140335	putative				
ORF135	0141754-0141077	porphobilinogen deaminase	U22968	*Yersinia pestis*	282	38
ORF136	0143141-0141780	unknown	D26185	*Bacillus subtilis*	1113	53
ORF137	0143829-0143128	ORF3	D64116	*Bacillus subtilis*	356	39
ORF138	0143923-0144393	putative				
ORF139	0144578-0146326	unknown	Z47210	*Streptococcus pneumoniae*	741	44
ORF140	0146413-0147078	manganese superoxide dismutase precursor	D12984	*Caenorhabditis elegans*	625	56
ORF141	0147140-0148075	acetyl-CoA carboxylase beta subunit (accD)	AE000604	*Helicobacter pylori*	704	52
ORF142	0148115-0148549	Dut	Z96072	*Mycobacterium tuberculosis*	277	53
ORF143	0148554-0149027	enzyme IIANtr	U18997	*Escherichia coli*	168	44

ORF	Position	Homology	ID	*Species*	Score	I%
ORF144	0149000-0149305	putative				
ORF145	0149229-0149708	enzyme IIANtr	U18997	*Escherichia coli*	169	43
ORF146	0149712-0150911	putative				
ORF147	0152044-0151004	putative				
ORF148	0152664-0151999	putative				
ORF149	0152900-0153352	hypothetical	U32702	*Haemophilus influenzae*	292	47
ORF150	0153389-0153997	hypothetical protein in purB 5' region	AE000213	*Escherichia coli*	555	49
ORF151	0155276-0153984	ClpC adenosine triphosphatase	U02604	*Bacillus subtilis*	986	45
ORF152	0156544-0155231	ClpC adenosine triphosphatase	U02604	*Bacillus subtilis*	1535	53
ORF153	0156806-0157525	putative				
ORF154	0157489-0158955	Unknown	Y08559	*Bacillus subtilis*	99	39
ORF155	0159104-0159961	putative				
ORF156	0159916-0161220	putative				
ORF157	0161183-0161593	glycine cleavage protein homolog	U12980	*Saccharomyces cerevisiae*	175	35
ORF158	0162333-0161623	unidentified protein of Na+-translocating NADH-quinone reductase	D49364	*Vibrio alginolyticus*	524	51
ORF159	0163001-0162363	NADH:uniquinone oxidoreductase	Z37111	*Vibrio alginolyticus*	543	55
ORF160	0163785-0162994	NADH:ubiquinone oxidoreductase (GP:Z37111_4)	U32702	*Haemophilus influenzae*	287	54
ORF161	0165499-0164474	NADH:ubiquinone oxidoreductase subunit B	Z37111	*Vibrio alginolyticus*	449	45
ORF162	0166482-0166093	H. pylori predicted coding region HP1542	AE000652	*Helicobacter pylori*	111	33
ORF163	0168093-0166729	pot. ORF 446 (aa 1-446)	X02369	*Bacillus subtilis*	722	42
ORF164	0169249-0168848	putative				
ORF165	0169586-0170431	hypothetical protein	D90906	*Synechocystis sp.*	462	48

ORF	Position	Homology	ID	*Species*	Score	I%
ORF166	0170780-0171334	putative				
ORF167	0171333-0172376	penicillin-binding protein 2	M26645	*Neisseria flavescens*	210	47
ORF168	0172309-0172722	penicillin-binding protein 2	M26645	*Neisseria flavescens*	176	44
ORF169	0173048-0174496	murE gene product	Z15056	*Bacillus subtilis*	789	43
ORF170	0174399-0174968	N-acetylmuramoyl-L-alanine amidase (amiA)	AE000589	*Helicobacter pylori*	177	41
ORF171	0175411-0175710	integration host factor beta subunit	L35259	*Pseudomonas aeruginosa*	110	38
ORF172	0175714-0177009	putative				
ORF173	0177423-0178115	carboxyltransferase alpha subunit	U59236	*Synechococcus PCC7942*	558	50
ORF174	0178240-0180021	ATP dependent translocator homolog (msbA)	U32691	*Haemophilus influenzae*	453	41
ORF175	0180704-0180048	putative				
ORF176	0181398-0180631	H. pylori predicted coding region HP0152	AE000536	*Helicobacter pylori*	256	34
ORF177	0182414-0181398	contains similarity to DNA polymerase III, alpha chain (SP:P47277)	AF007270	*Arabidopsis thaliana*	173	50
ORF178	0182913-0183656	putative Ptc1 protein	Y13937	*Bacillus subtilis*	371	53
ORF179	0183665-0184786	NifS2	AF008220	*Bacillus subtilis*	452	43
ORF180	0185962-0184796	similar to [SwissProt Accession Number P37908]	D90888	*Escherichia coli*	93	30
ORF181	0186848-0186000	hypothetical	U32728	*Haemophilus influenzae*	154	35
ORF182	0187270-0186749	putative				
ORF183	0187426-0187809	regulatory protein for beta-lactamase	D90902	*Synechocystis sp.*	96	36
ORF184	0189481-0188798	putative				
ORF185	0189693-0190352	prolipoprotein diacylglyceryl transferase	AJ000977	*Rhodobacter sphaeroides*	99	38
ORF186	0190235-0190510	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF187	0190785-0191786	putative				
ORF188	0191790-0192464	putative				
ORF189	0192392-0193183	60 kDa inner-membrane protein	AE000645	*Helicobacter pylori*	373	40
ORF190	0193254-0194630	DnaA	D89066	*Staphylococcus aureus*	545	43
ORF191	0195046-0194690	putative				
ORF192	0195184-0197031	glycogen phosphorylase B	U47025	*Homo sapiens*	1758	56
ORF193	0197018-0197635	glycogen phosphorylase (AA 1 790)	X16931	*Escherichia coli*	580	53
ORF194	0197762-0198208	unknown	X86470	*Saccharomyces cerevisiae*	148	42
ORF195	0198963-0197668	F23B12.5	Z77659	*Caenorhabditis elegans*	795	50
ORF196	0199957-0198962	pyruvate dehydrogenase E1 beta subunit	U09137	*Arabidopsis thaliana*	856	48
ORF197	0200327-0199941	pyruvate dehydrogenase E1 component, alpha subunit	U38804	*Porphyra purpurea*	170	31
ORF198	0200685-0200266	pyruvate dehydrogenase complex E1 alpha subunit	U81808	*Thiobacillus ferrooxidans*	302	60
ORF199	0200896-0200585	TPP-dependent acetoin dehydrogenase alpha-subunit	L31844	*Clostridium magnum*	127	43
ORF200	0201169-0202377	putative				
ORF201	0203441-0202380	UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	U70214	*Escherichia coli*	577	38
ORF202	0203998-0203471	putative				
ORF203	0206401-0204059	OMP1 precursor	U51683	*Brucella abortus*	83	31
ORF204	0207425-0206811	recombination protein	D90916	*Synechocystis sp.*	334	40
ORF205	0207548-0208528	beta-ketoacyl-acyl carrier protein synthase III	M77744	*Escherichia coli*	706	50
ORF206	0208548-0209471	malonyl-CoA:Acyl carrier protein transacylase	U59433	*Bacillus subtilis*	522	48

ORF	Position	Homology	ID	*Species*	Score	I%
ORF207	0209471-0210214	3-ketoacyl-acyl carrier protein reductase	U59433	*Bacillus subtilis*	616	51
ORF208	0210586-0210816	acyl carrier protein (acpP)	U32701	*Haemophilus influenzae*	220	57
ORF209	0211332-0210883	protein kinase type II regulatory subunit (, EC 2.7.1.37)	J02934	*Rattus norvegicus*	150	31
ORF210	0212978-0211374	putative				
ORF211	0214080-0212875	unknown	AF017105	*Chlamydia psittaci*	852	63
ORF212	0214710-0214168	inclusion membrane protein C	AF017105	*Chlamydia psittaci*	231	43
ORF213	0215143-0214754	inclusion membrane protein B	AF017105	*Chlamydia psittaci*	181	47
ORF214	0216705-0215236	sodium-dependent transporter	AF017105	*Chlamydia psittaci*	1341	70
ORF215	0217917-0216892	amino acid transporter	AF017105	*Chlamydia psittaci*	1027	60
ORF216	0217088-0217441	putative				
ORF217	0218364-0218702	putative				
ORF218	0218695-0219009	putative				
ORF219	0219179-0219748	putative				
ORF220	0219891-0220430	putative				
ORF221	0220499-0221074	putative				
ORF222	0221137-0221541	putative				
ORF223	0221601-0222092	putative				
ORF224	0222472-0223290	putative				
ORF225	0223435-0223818	LAGLI-DADG endonuclease	U57090	*Chlamydia trachomatis*	619	99
ORF226	0224278-0225171	YqfU	D84432	*Bacillus subtilis*	530	46
ORF227	0225728-0225174	phenylacrylic acid decarboxylase	U67467	*Methanococcus jannaschii*	334	52
ORF228	0225334-0225549	Ydr537cp	U43834	*Saccharomyces cerevisiae*	96	42
ORF229	0226612-0225749	4-hydroxybenzoate octaprenyltransferase	U61168	*Bacillus firmus*	321	36

ORF	Position	Homology	ID	*Species*	Score	I%
ORF230	0227299-0226769	putative				
ORF231	0227616-0227161	stationary-phase survival protein (surE)	AE000602	*Helicobacter pylori*	274	48
ORF232	0228457-0227750	f311; This 311 aa ORF is 22 pct identical (13 gaps) to 186 residues of an approx. 488 aa protein YACA_BACSU SW: P37563; pyu1 of D21139	AE000232	*Escherichia coli*	246	36
ORF233	0230001-0228607	GadC	AF005098	*Lactococcus lactis*	740	35
ORF234	0231074-0230151	f374; This 374 aa ORF is 30 pct identical (9 gaps) to 102 residues of an approx. 512 aa protein FLIC_SALMU SW: P06177	AE000299	*Escherichia coli*	985	65
ORF235	0231348-0233006	putative				
ORF236	0233134-0233829	orf2	D88555	*Methanobacterium thermoautotrophicum*	351	52
ORF237	0233855-0234265	hypothetical protein	D90906	*Synechocystis sp.*	151	37
ORF238	0234282-0234854	ORF_o211	U28377	*Escherichia coli*	105	54
ORF239	0236117-0235227	glutamate 1-semialdehyde 2,1-aminomutase	X82072	*Pseudomonas aeruginosa*	650	52
ORF240	0236314-0238209	leucine tRNA synthetase	AF008220	*Bacillus subtilis*	1836	61
ORF241	0238164-0238769	leucine tRNA synthetase	AF008220	*Bacillus subtilis*	410	46
ORF242	0238769-0240061	3-deoxy-D-manno-2-octulosonic acid (Kdo) transferase	Z22659	*Chlamydia trachomatis*	2240	100
ORF243	0241980-0240313	pyrophosphate-dependent phosphofructokinase beta subunit	Z32850	*Ricinus communis*	1021	43
ORF244	0242846-0241941	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF245	0244480-0242798	pyrophosphate-dependent phosphofructokinase beta subunit	Z32850	*Ricinus communis*	1017	42
ORF246	0245897-0244479	YflS	D86417	*Bacillus subtilis*	951	42
ORF247	0246877-0245924	putative				
ORF248	0247731-0246985	ATP binding protein	L18760	*Lactococcus lactis*	442	47
ORF249	0248585-0247743	sporulation protein	M57689	*Bacillus subtilis*	532	38
ORF250	0249420-0248569	sporulation protein	M57689	*Bacillus subtilis*	601	38
ORF251	0250383-0249766	sporulation protein	M57689	*Bacillus subtilis*	464	47
ORF252	0251186-0250545	oligopeptide permease homolog AII	AF000366	*Borrelia burgdorferi*	119	31
ORF253	0252111-0251095	sporulation protein	M57689	*Bacillus subtilis*	317	36
ORF254	0253088-0252066	P. haemolytica o-sialoglycoprotein endopeptidase; P36175 (660) transmembrane	D88802	*Bacillus subtilis*	601	46
ORF255	0255234-0256718	Mg2+ transporter	D90905	*Synechocystis sp.*	103	35
ORF256	0256762-0257844	tRNA guanine transglycosylase	L33777	*Zymomonas mobilis*	482	44
ORF257	0257911-0258690	putative				
ORF258	0258780-0259187	putative				
ORF259	0259193-0261604	subunit B of DNA gyrase	Y07916	*Salmonella typhimurium*	1925	58
ORF260	0261622-0264129	DNA gyrase	L47978	*Aeromonas salmonicida*	1963	45
ORF261	0264125-0264742	unknown	D26185	*Bacillus subtilis*	307	37
ORF262	0264741-0265628	replication protein (dnaX)	U32802	*Haemophilus influenzae*	162	35
ORF263	0266362-0265631	putative isozyme of glucose-6-P-dehydrogenase; developmentally regulated gene in heterocyst development	U14553	*Anabaena sp.*	218	47
ORF264	0266938-0266426	glucose 6-phosphate dehydrogenase	U83195	*Chlamydia trachomatis*	914	99

ORF	Position	Homology	ID	*Species*	Score	I%
ORF265	0267961-0266942	glucose 6-phosphate dehydrogenase	U83195	*Chlamydia trachomatis*	1770	99
ORF266	0268320-0268066	ORF3	U15192	*Chlamydia trachomatis*	403	100
ORF267	0268510-0268205	ORF3	U15192	*Chlamydia trachomatis*	320	91
ORF268	0270116-0268500	CTP synthetase	U15192	*Chlamydia trachomatis*	2828	100
ORF269	0270856-0270095	CMP-2-keto-3-deoxyoctulosonic acid synthetase	U15192	*Chlamydia trachomatis*	1313	100
ORF270	0271191-0271613	putative				
ORF271	0272219-0272932	nitrate transporter	X61625	*Synechococcus sp.*	300	34
ORF272	0272884-0273588	putative				
ORF273	0274816-0273596	putative				
ORF274	0274821-0275666	putative				
ORF275	0277689-0276103	ORF_f535	U28377	*Escherichia coli*	396	38
ORF276	0278268-0278816	putative				
ORF277	0279771-0279013	tryptophan synthase alpha subunit	M15826	*Pseudomonas aeruginosa*	357	37
ORF278	0280777-0279767	tryptophan synthetase	M91661	*Coprinus cinereus*	1042	62
ORF279	0281603-0281295	tryptophan repressor	L26582	*Enterobacter aerogenes*	151	35
ORF280	0282104-0281787	putative				
ORF281	0284335-0282794	putative				
ORF282	0284460-0284795	putative				
ORF283	0284817-0285674	putative				
ORF284	0285637-0286137	putative				
ORF285	0286357-0286677	putative				
ORF286	0286852-0287898	hypothetical protein	U88070	*Chlamydia psittaci*	99	35
ORF287	0288127-0289227	comE ORF1	D64002	*Synechocystis sp.*	90	46
ORF288	0289744-0290679	hypothetical protein	U88070	*Chlamydia psittaci*	246	36
ORF289	0290828-0291535	putative				
ORF290	0291604-0292230	endonuclease	U09868	*Escherichia coli*	160	37
ORF291	0292326-0293048	putative				
ORF292	0293330-0294853	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF293	0295684-0295010	glutamine transport ATP-binding protein Q	U67524	*Methanococcus jannaschii*	407	38
ORF294	0296336-0295692	H. influenzae predicted coding region HI1555	U32830	*Haemophilus influenzae*	134	37
ORF295	0297238-0296243	putative				
ORF296	0297791-0298735	putative				
ORF297	0298905-0300458	similar to putative oxygenase of S. fradiae	U73857	*Escherichia coli*	82	40
ORF298	0302152-0300527	putative				
ORF299	0304917-0302071	putative				
ORF300	0306157-0304973	DNA ligase	M74792	*Thermus aquaticus thermophilus*	745	41
ORF301	0306494-0306111	DNA LIGASE (EC 6.5.1.2) (POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+)).	D90870	*Escherichia coli*	197	40
ORF302	0306963-0306436	Mycoplasma pneumoniae, DNA ligase; similar to Swiss-Prot Accession Number P15042, from E. coli	AE000047	*Mycoplasma pneumoniae*	292	37
ORF303	0308773-0306977	unknown	Z84395	*Mycobacterium tuberculosis*	316	52
ORF304	0309881-0309276	putative				
ORF305	0310720-0309872	putative				
ORF306	0311570-0310716	putative				
ORF307	0312451-0311972	Preprotein translocase SecA subunit.	D90832	*Escherichia coli*	123	86
ORF308	0313435-0314364	sporulation protein	M57689	*Bacillus subtilis*	202	37
ORF309	0314340-0314738	putative				
ORF310	0315526-0314741	orfX gene product	X58778	*Klebsiella pneumoniae*	169	45
ORF311	0316507-0315665	Similar to Saccharomyces cerevisiae SUA5 protein	Z38002	*Bacillus subtilis*	147	41

ORF	Position	Homology	ID	*Species*	Score	I%
ORF312	0317224-0316529	serine esterase [Spirulina platensis, C1, Peptide, 207 aa]	S70419	*Spirulina platensis*	167	58
ORF313	0317592-0317338	putative				
ORF314	0318470-0317499	putative				
ORF315	0317599-0317874	putative				
ORF316	0318947-0318477	putative				
ORF317	0319342-0320142	ORF2	L35036	*Chlamydia psittaci*	802	60
ORF318	0320544-0321497	putative				
ORF319	0321485-0321937	putative				
ORF320	0321901-0322362	putative				
ORF321	0322301-0323140	putative				
ORF322	0323144-0324913	putative				
ORF323	0325621-0324977	YqiZ	D84432	*Bacillus subtilis*	430	43
ORF324	0326268-0325621	integral membrane protein homolog	U97348	*Lactobacillus fermentum*	343	44
ORF325	0326469-0327203	adenylate kinase	AB000111	*Synechococcus sp.*	371	46
ORF326	0327281-0328150	putative				
ORF327	0328605-0328204	RpsI	Z95389	*Mycobacterium tuberculosis*	315	55
ORF328	0329045-0328734	50S ribosomal subunit protein L13	U18997	*Escherichia coli*	269	60
ORF329	0329663-0329292	YqhX	D84432	*Bacillus subtilis*	297	56
ORF330	0330666-0329608	biotin carboxylase	L14862	*Anabaena sp.*	1089	58
ORF331	0331161-0330670	YqhW	D84432	*Bacillus subtilis*	208	52
ORF332	0331731-0331177	elongation factor P	D64001	*Synechocystis sp.*	297	33
ORF333	0332404-0331721	putative CfxE protein	Y13937	*Bacillus subtilis*	483	55
ORF334	0332779-0333021	putative				
ORF335	0333005-0333589	putative				
ORF336	0334357-0333806	putative				
ORF337	0334089-0334361	putative				
ORF338	0335142-0334729	putative				
ORF339	0335195-0335602	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF340	0335673-0335194	putative				
ORF341	0336334-0335903	putative				
ORF342	0337378-0336338	putative				
ORF343	0339947-0337347	ATP-dependent protease binding subunit	M29364	*Escherichia coli*	2005	53
ORF344	0340507-0341847	Pz-peptidase	D88209	*Bacillus licheniformis*	508	39
ORF345	0341783-0342022	group B oligopeptidase PepB	U49821	*Streptococcus agalactiae*	140	48
ORF346	0342249-0342470	hypA protein	M31739	*Chlamydia trachomatis*	361	99
ORF347	0342597-0343370	heat shock protein	L12004	*Chlamydia trachomatis*	1271	99
ORF348	0343361-0344032	hypB protein	M31739	*Chlamydia trachomatis*	1051	100
ORF349	0343956-0344225	hypB protein	M31739	*Chlamydia trachomatis*	344	100
ORF350	0344357-0345142	orf 3'of chaperonin homolog hypB [Chlamydia psittaci, pigeon strain P-1041, Peptide Partial, 98 aa]	S40172	*Chlamydia psittaci*	344	63
ORF351	0345913-0345161	o247; This 247 aa ORF is 51 pct identical (0 gaps) to 117 residues of an approx. 160 aa protein YPH7_CHRVI SW: P45371	AE000174	*Escherichia coli*	387	41
ORF352	0347102-0346080	mutY homolog	U63329	*Homo sapiens*	492	46
ORF353	0347113-0347940	hypothetical 36.0 kD protein in rne-rpmF intergenic region	AE000209	*Escherichia coli*	397	44
ORF354	0350164-0348146	putative				
ORF355	0350423-0351283	enoyl-acyl carrier protein reductase [Brassica napus, Peptide, 385 aa]	S60064	*Brassica napus*	909	64
ORF356	0352207-0351314	hypothetical protein	D90914	*Synechocystis sp.*	113	42
ORF357	0352727-0352245	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF358	0353661-0353305	FUNCTION UNKNOWN, SIMILAR PRODUCT IN E. COLI AND MYCOPLASMA PNEUMONIAE.	AB001488	*Bacillus subtilis*	213	40
ORF359	0354218-0353670	NADPH thioredoxin reductase	Z23108	*Arabidopsis thaliana*	577	60
ORF360	0354604-0354140	Thioredoxin Reductase (NADPH)	D45049	*Neurospora crassa*	417	60
ORF361	0355059-0356672	30S ribosomal protein S1	D90729	*Escherichia coli*	1305	44
ORF362	0356793-0357377	NusA	U74759	*Chlamydia trachomatis*	948	100
ORF363	0357326-0358093	NusA	U74759	*Chlamydia trachomatis*	1216	100
ORF364	0358053-0360743		U74759	*Chlamydia trachomatis*	3311	98
ORF365	0360753-0361121	ORF6 gene product	Z18631	*Bacillus subtilis*	116	32
ORF366	0361183-0361884	tRNA pseudouridine 55 synthase	D90917	*Synechocystis sp.*	362	42
ORF367	0361826-0362746	protein X	M35367	*Pseudomonas fluorescens*	192	49
ORF368	0363853-0362816	hypothetical GTP-binding protein in pth 3' region	AE000219	*Escherichia coli*	978	52
ORF369	0364116-0365195	cds1 gene product	U88070	*Chlamydia psittaci*	1631	88
ORF370	0365198-0365587	cds2 gene product	U88070	*Chlamydia psittaci*	516	93
ORF371	0365479-0367320	cds2 gene product	U88070	*Chlamydia psittaci*	2817	87
ORF372	0367341-0368603	copN gene product	U88070	*Chlamydia psittaci*	585	37
ORF373	0368644-0369081	scc1 gene product	U88070	*Chlamydia psittaci*	528	67
ORF374	0369088-0370251	No definition line found	U88070	*Chlamydia psittaci*	1362	62
ORF375	0370820-0371086	ribosomal protein L28 (rpL28)	U32776	*Haemophilus influenzae*	182	46
ORF376	0371203-0372816	hypothetical protein	U88070	*Chlamydia psittaci*	1926	68
ORF377	0373119-0373529	hypothetical protein	U88070	*Chlamydia psittaci*	286	49
ORF378	0373614-0374204	hypothetical protein	U88070	*Chlamydia psittaci*	379	48
ORF379	0374736-0374224	putative				
ORF380	0376391-0374703	putative				
ORF381	0377038-0376748	corresponds to a 97 amino acid long polypeptide	L40838	*Chlamydia trachomatis*	490	98

ORF	Position	Homology	ID	*Species*	Score	I%
ORF382	0377853-0378737	methylenetetrahydrofolate dehydrogenase	D64000	*Synechocystis sp.*	678	51
ORF383	0378626-0379048	putative				
ORF384	0379017-0379403	hypothetical	U32702	*Haemophilus influenzae*	137	45
ORF385	0380009-0379641	small protein	D90914	*Synechocystis sp.*	216	51
ORF386	0380373-0381470	DNA polymerase III beta-subunit (dnaN)	U32780	*Haemophilus influenzae*	76	39
ORF387	0381473-0382567	recombination protein	D26185	*Bacillus subtilis*	477	35
ORF388	0382704-0383702	putative				
ORF389	0383945-0383655	hypothetical	U70214	*Escherichia coli*	134	35
ORF390	0385217-0383949	putative				
ORF391	0385507-0385178	conserved hypothetical secreted protein	AE000606	*Helicobacter pylori*	185	45
ORF392	0386845-0385706	hypothetical protein	D64000	*Synechocystis sp.*	686	41
ORF393	0386127-0386627	putative				
ORF394	0387372-0386872	ORF1; putative	M26130	*Streptococcus parasanguis*	150	35
ORF395	0387724-0387338	ytgD	AF008220	*Bacillus subtilis*	168	42
ORF396	0388250-0387816	TroR	U55214	*Treponema pallidum*	134	40
ORF397	0389169-0388237	putative protein of 299 amino acids	U30821	*Cyanophora paradoxa*	164	31
ORF398	0389955-0389173	TroB	U55214	*Treponema pallidum*	592	51
ORF399	0390895-0389945	YtgA	AF008220	*Bacillus subtilis*	282	30
ORF400	0391514-0391810	putative				
ORF401	0392410-0393996	adenine nucleotide translocase	Z49227	*Arabidopsis thaliana*	1295	56
ORF402	0394170-0395354	lepA gene product	X91655	*Bacillus subtilis*	1235	60
ORF403	0395309-0395992	GTP-binding membrane protein (lepA)	AE000552	*Helicobacter pylori*	543	54
ORF404	0396538-0396059	phosphogluconate dehydrogenase	U30255	*Homo sapiens*	411	55
ORF405	0397507-0396542	6-phosphogluconate dehydrogenase	AB006102	*Candida albicans*	908	51

ORF	Position	Homology	ID	*Species*	Score	I%
ORF406	0398753-0397401	tyrosyl-tRNA synthetase	M13148	*Bacillus caldotenax*	844	45
ORF407	0399688-0398909	whiG-Stv gene product	X68709	*Streptoverticillium griseocarneum*	463	41
ORF408	0400167-0399778	FLHA gene product	X63698	*Bacillus subtilis*	134	35
ORF409	0401224-0400034	flbF	M73782	*Caulobacter crescentus*	355	39
ORF410	0401776-0402021	ferredoxin IV	M59855	*Rhodobacter capsulatus*	98	54
ORF411	0402126-0403220	putative				
ORF412	0403348-0405180	GcpE	D90908	*Synechocystis sp.*	995	49
ORF413	0403788-0403276	putative				
ORF414	0405165-0405920	YfiH	U50134	*Escherichia coli*	166	43
ORF415	0407049-0405955	dihydrolipoamide transsuccinylase (odhB; EC 2.3.1.61)	M27141	*Bacillus subtilis*	833	61
ORF416	0409773-0407056	alpha-ketoglutarate dehydrogenase	U41762	*Rhodobacter capsulatus*	1537	50
ORF417	0410532-0411416	YqeR	D84432	*Bacillus subtilis*	496	44
ORF418	0411707-0413410	putative				
ORF419	0413433-0412606	putative				
ORF420	0413404-0413952	putative				
ORF421	0413841-0415112	putative				
ORF422	0414379-0413978	putative				
ORF423	0416664-0415177	putative				
ORF424	0417450-0416740	unknown	Z94752	*Mycobacterium tuberculosis*	172	36
ORF425	0418053-0417721	putative				
ORF426	0418603-0418031	putative				
ORF427	0419525-0418647	Hc2 nucleoprotein	L10193	*Chlamydia trachomatis*	1661	92
ORF428	0420037-0419672	[karp] gene products	M86605	*Chlamydia trachomatis*	612	96
ORF429	0421078-0420245	aminopeptidase	D17450	*Mycoplasma salivarium*	269	41
ORF430	0421988-0421518	putative	L39923	*Mycobacterium leprae*	165	36
ORF431	0422486-0423043	putative				
ORF432	0423226-0425079	glycogen operon protein GlgX	D90908	*Synechocystis sp.*	1229	55

ORF	Position	Homology	ID	*Species*	Score	I%
ORF433	0426054-0425146	putative				
ORF434	0426985-0426245	Holliday junction specific DNA helicase	D83138	*Pseudomonas aeruginosa*	633	53
ORF435	0427248-0427817	deoxycytidine triphosphate deaminase (dcd)	AE000554	*Helicobacter pylori*	612	63
ORF436	0429560-0429886	putative				
ORF437	0430360-0429857	biotin apo-protein ligase	U27182	*Saccharomyces cerevisiae*	173	38
ORF438	0430637-0430323	putative				
ORF439	0430933-0431787	putative				
ORF440	0431658-0431987	putative				
ORF441	0432232-0434475	exonuclease V alpha-subunit	U29581	*Escherichia coli*	289	53
ORF442	0436308-0434620	methionyl-tRNA synthetase	AB004537	*Schizosaccharomyces pombe*	817	54
ORF443	0436574-0436272	putative				
ORF444	0437685-0436567	RNAseH II	AF005098	*Lactococcus lactis*	395	47
ORF445	0438262-0437894	ribosomal protein L19	X72627	*Synechocystis sp.*	287	47
ORF446	0439127-0438285	tRNA (guanine-N1)-methyltransferase (trmD)	U32705	*Haemophilus influenzae*	374	56
ORF447	0439339-0438986	tRNA (guanine-N1)-methyltransferase (trmD)	U32705	*Haemophilus influenzae*	199	57
ORF448	0439702-0439358	ribosomal protein S16 (rpS16)	U32705	*Haemophilus influenzae*	168	39
ORF449	0441042-0439699	signal recognition particle protein	AE000347	*Escherichia coli*	865	40
ORF450	0441911-0441042	product similar to E.coli PRFA2 protein	Z49782	*Bacillus subtilis*	314	37
ORF451	0442593-0441898	polypeptide chain release factor 1 (prfA)	U32830	*Haemophilus influenzae*	708	62
ORF452	0444688-0446388	leader peptidase I	D90904	*Synechocystis sp.*	268	44
ORF453	0448068-0446452	isoleucyl-tRNA synthetase	U04953	*Homo sapiens*	704	49
ORF454	0449560-0447932	isoleucyl-tRNA synthetase	U04953	*Homo sapiens*	1687	55

ORF	Position	Homology	ID	*Species*	Score	I%
ORF455	0450546-0451076	putative				
ORF456	0451623-0451144	putative				
ORF457	0452593-0451517	putative				
ORF458	0453195-0452632	putative				
ORF459	0453567-0454868	product similar to E. coli PhoH protein	Z97025	*Bacillus subtilis*	820	50
ORF460	0455430-0454972	CydB	Z95554	*Mycobacterium tuberculosis*	105	31
ORF461	0456032-0455367	cyanide insensitive terminal oxidase	Y10528	*Pseudomonas aeruginosa*	388	38
ORF462	0457384-0456047	cyanide insensitive terminal oxidase	Y10528	*Pseudomonas aeruginosa*	537	52
ORF463	0457659-0458450	YbbP	AB002150	*Bacillus subtilis*	324	42
ORF464	0458508-0459632	putative				
ORF465	0459839-0461203	HtrB protein	X61000	*Escherichia coli*	77	31
ORF466	0461624-0461196	unknown	U87792	*Bacillus subtilis*	114	38
ORF467	0461887-0462621	hypothetical protein	Z75208	*Bacillus subtilis*	148	51
ORF468	0463758-0462895	putative				
ORF469	0464048-0464629	putative				
ORF470	0464721-0465848	putative				
ORF471	0467420-0466113	PET112	D90913	*Synechocystis sp.*	892	48
ORF472	0468891-0467419	amidase	U49269	*Moraxella catarrhalis*	1051	46
ORF473	0469280-0468906	putative				
ORF474	0469349-0469675	putative				
ORF475	0471226-0469826	putative				
ORF476	0471624-0471106	putative				
ORF477	0471954-0473267	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	173	33
ORF478	0473252-0473695	POMP90A precursor	U65942	*Chlamydia psittaci*	175	39
ORF479	0473982-0474527	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	193	38
ORF480	0475198-0474602	putative				
ORF481	0476527-0475613	POMP91A	U65942	*Chlamydia psittaci*	100	38

ORF	Position	Homology	ID	*Species*	Score	I%
ORF482	0478640-0476517	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	537	40
ORF483	0479084-0478665	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	234	35
ORF484	0479723-0479088	putative outer membrane protein	U72499	*Chlamydia psittaci*	313	40
ORF485	0480012-0479668	putative				
ORF486	0481466-0479895	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	391	38
ORF487	0481732-0481496	putative				
ORF488	0481864-0483429	POMP90A precursor	U65942	*Chlamydia psittaci*	114	40
ORF489	0483402-0484964	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	77	34
ORF490	0484898-0487864	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	506	39
ORF491	0485725-0485222	putative				
ORF492	0488204-0489247	putative				
ORF493	0488571-0488233	putative				
ORF494	0489440-0490456	putative				
ORF495	0492765-0490507	branching enzyme	M31544	*Synechococcus PCC6301*	1624	57
ORF496	0492357-0492893	putative				
ORF497	0493744-0492737	putative				
ORF498	0493971-0494675	YqkM	D84432	*Bacillus subtilis*	230	44
ORF499	0494573-0494869	xprB	M54884	*Escherichia coli*	245	48
ORF500	0494835-0495365	putative				
ORF501	0495174-0494872	putative				
ORF502	0495687-0496634	putative				
ORF503	0496295-0497176	putative				
ORF504	0497703-0498515	putative				
ORF505	0498280-0499239	putative				
ORF506	0499215-0500732	putative				
ORF507	0501710-0500790	penicillin tolerance protein (lytB)	U32781	*Haemophilus influenzae*	702	50

ORF	Position	Homology	ID	*Species*	Score	I%
ORF508	0502863-0501808	putative				
ORF509	0503675-0502692	putative				
ORF510	0504984-0503722	hypothetical protein	Z96072	*Mycobacterium tuberculosis*	102	42
ORF511	0505763-0506986	hypothetical protein in pth-prs intergenic region	AE000219	*Escherichia coli*	740	44
ORF512	0506999-0507439	putative				
ORF513	0508404-0507649	fumarate hydratase	AF013216	*Myxococcus xanthus*	611	54
ORF514	0508291-0508590	putative				
ORF515	0508915-0508478	fumarase	D64000	*Synechocystis sp.*	386	57
ORF516	0509753-0510691	thiamine-repressed protein (nmt1)	U32720	*Haemophilus influenzae*	82	31
ORF517	0511039-0511527	putative				
ORF518	0511547-0512185	hypothetical protein (SP:P46851)	U67608	*Methanococcus jannaschii*	208	39
ORF519	0512382-0513092	methionine amino peptidase	M15106	*Escherichia coli*	384	46
ORF520	0514287-0513055	putative				
ORF521	0514789-0515244	putative				
ORF522	0514994-0515269	putative				
ORF523	0515553-0515804	putative				
ORF524	0515808-0516422	putative				
ORF525	0516476-0517171	putative				
ORF526	0517888-0517400	orf150 gene product	X95938	*Porphyromonas gingivalis*	340	51
ORF527	0518114-0518380	30S ribosomal protein S15	D90901	*Synechocystis sp.*	245	52
ORF528	0518403-0518822	polynucleotide phosphorylase	AF010578	*Pisum sativum*	306	49
ORF529	0518923-0519516	polyribonucleotide phophorylase	U52048	*Spinacia oleracea*	387	47
ORF530	0519577-0520497	polynucleotide phosphorylase	U18997	*Escherichia coli*	860	54
ORF531	0521986-0520718	ATP-binding protein	U01376	*Escherichia coli*	970	49

ORF	Position	Homology	ID	*Species*	Score	I%
ORF532	0522131-0521886	cell division protein (ftsH)	U32812	*Haemophilus influenzae*	314	76
ORF533	0523495-0522143	putative				
ORF534	0524591-0523623	ORF327 gene product	U38804	*Porphyra purpurea*	148	44
ORF535	0524652-0525746	putative				
ORF536	0525731-0526078	putative				
ORF537	0525939-0526400	putative				
ORF538	0526301-0526735	putative				
ORF539	0528323-0526851	putative				
ORF540	0528861-0528292	putative				
ORF541	0529723-0529142	phenylalanyl-tRNA synthetase alpha subunit	X53057	*Bacillus subtilis*	476	52
ORF542	0530166-0529624	phenylalany-tRNA synthetase beta subunit	Z75208	*Bacillus subtilis*	164	40
ORF543	0530543-0530223	ribosomal protein L20 (AA 1-119)	X16188	*Bacillus stearothermophilus*	230	47
ORF544	0531378-0530737	unknown	Z85982	*Mycobacterium tuberculosis*	452	50
ORF545	0532370-0533272	UDP-N-acetylenolpyruvylglucosamine reductase	U86147	*Synechococcus PCC7942*	488	43
ORF546	0533804-0533244	YtqB	AF008220	*Bacillus subtilis*	273	38
ORF547	0534672-0533944	hypothetical protein MTCY08D5.03c	Z92669	*Mycobacterium tuberculosis*	170	35
ORF548	0535915-0534878	ribonucleoside diphosphate reductase, beta subunit (nrdB)	AE000553	*Helicobacter pylori*	397	33
ORF549	0539153-0535956	ribonucleoside-diphosphate reductase 1 alpha subunit (nrdA)	AE000581	*Helicobacter pylori*	1447	51
ORF550	0539779-0540519	phosphatidylserine synthase (pssA)	AE000614	*Helicobacter pylori*	226	49
ORF551	0540523-0540969	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF552	0540906-0541805	hypothetical 54.7 kD protein in udp 3' region precursor (o475)	AE000459	*Escherichia coli*	82	39
ORF553	0543255-0541825	Ydr430cp; CAI: 0.15	U33007	*Saccharomyces cerevisiae*	130	48
ORF554	0544133-0543222	putative				
ORF555	0544565-0544179	hypA gene product	X86493	*Clostridium perfringens*	221	46
ORF556	0544762-0544487	orf1 gene product	X70951	*Saccharomyces cerevisiae*	153	38
ORF557	0546270-0544951	serine protease (htrA)	AE000610	*Helicobacter pylori*	981	46
ORF558	0547480-0546584	succinyl coenzyme A synthetase alpha subunit	U23408	*Dictyostelium discoideum*	869	63
ORF559	0546789-0547382	putative				
ORF560	0547901-0547476	putative succinyl-coA synthetase beta chain	AJ000975	*Bacillus subtilis*	388	55
ORF561	0548634-0547900	succinate--CoA ligase (ADP-forming)	X54073	*Thermus aquaticus flavus*	498	46
ORF562	0548692-0549459	cell division protein (ftsY)	AE000588	*Helicobacter pylori*	330	46
ORF563	0550385-0549663	putative				
ORF564	0551605-0550421	Tyrosine-specific transport protein (Tyrosine permease).	D90832	*Escherichia coli*	508	40
ORF565	0552990-0551797	tyrosine-specific transport protein (tyrP)	U32730	*Haemophilus influenzae*	353	36
ORF566	0554913-0553096	L-glutamine:D-fructose-6-P amidotransferase precursor	U17352	*Thermus aquaticus thermophilus*	1324	45
ORF567	0556300-0554927	hypothetical	U32824	*Haemophilus influenzae*	1009	51
ORF568	0556524-0556904	putative				
ORF569	0558126-0557314	putative				
ORF570	0557810-0558235	putative				
ORF571	0559215-0558310	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF572	0561349-0559196	POMP91A	U65942	*Chlamydia psittaci*	245	39
ORF573	0562931-0561150	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	130	38
ORF574	0564083-0563121	putative PlsX protein	Y13937	*Bacillus subtilis*	519	45
ORF575	0563593-0563943	putative				
ORF576	0565400-0566953	ORF_f495; orfF of ECMRED, uses 2nd start	U18997	*Escherichia coli*	874	39
ORF577	0567079-0567966	glycerol-3-phosphate acyltransferase	M80571	*Cucumis sativus*	594	45
ORF578	0568093-0570399	insulin-degrading enzyme	M58465	*Drosophila melanogaster*	334	42
ORF579	0571269-0572021	putative				
ORF580	0572519-0572755	putative				
ORF581	0573519-0572731	unknown	Z94752	*Mycobacterium tuberculosis*	203	35
ORF582	0572879-0573427	putative				
ORF583	0573890-0573660	putative heat shock protein ORF; putative	M62820	*Chlamydia trachomatis*	315	83
ORF584	0574426-0574184	ribosomal protein S18 homolog; putative	M62820	*Chlamydia trachomatis*	384	99
ORF585	0574781-0574446	ribosomal protein S6 (rps6)	AE000630	*Helicobacter pylori*	176	39
ORF586	0575243-0574923	peptidyl-tRNA hydrolase	U31570	*Chlamydia trachomatis*	358	78
ORF587	0575458-0575057	peptidyl-tRNA hydrolase	U31570	*Chlamydia trachomatis*	393	81
ORF588	0575849-0575469	partial ctc gene product (AA 1-186)	X16518	*Bacillus subtilis*	94	37
ORF589	0576602-0578023	glycogen (starch) synthase	D90899	*Synechocystis sp.*	695	48
ORF590	0578640-0578017	phosphatidylglycerophosphate synthase	U87792	*Bacillus subtilis*	243	48
ORF591	0579096-0582104	glycyl-tRNA synthetase	U20547	*Chlamydia trachomatis*	5054	99
ORF592	0582697-0582206	putative				
ORF593	0583122-0582811	putative				
ORF594	0583514-0583182	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF595	0583869-0583438	putative				
ORF596	0584435-0583827	dnaG	AB001896	*Staphylococcus aureus*	298	41
ORF597	0584967-0584299	DNA primase	U13165	*Listeria monocytogenes*	339	41
ORF598	0585297-0585016	putative				
ORF599	0585240-0586610	DNA mismatch repair protein	D90909	*Synechocystis sp.*	673	42
ORF600	0586484-0587758	DNA mismatch repair protein	U71154	*Aquifex pyrophilus*	845	50
ORF601	0587786-0589408	excinuclease ABC subunit C (uvrC)	U32691	*Haemophilus influenzae*	719	46
ORF602	0589198-0589578	exinuclease ABC subunit C	U29587	*Rhodobacter sphaeroides*	156	42
ORF603	0590061-0589630	putative				
ORF604	0590739-0591272	putative				
ORF605	0592412-0592765	homologous to E.coli rnpA	X62539	*Bacillus subtilis*	117	34
ORF606	0593145-0592849	putative				
ORF607	0593900-0593121	putative				
ORF608	0594195-0595637	cys-tRNA synthetase (cysS)	U32693	*Haemophilus influenzae*	991	49
ORF609	0596122-0595640	lysyl-tRNA synthetase	D90906	*Synechocystis sp.*	375	53
ORF610	0596864-0596154	lysine--tRNA ligase	X70708	*Thermus aquaticus thermophilus*	571	52
ORF611	0597731-0597282	putative				
ORF612	0598524-0600809	putative PriA protein	Y13937	*Bacillus subtilis*	1097	38
ORF613	0601876-0600734	L-alanine - pimelyl CoA ligase	U51868	*Bacillus subtilis*	242	42
ORF614	0603523-0601910	2-acylglycerophospho-ethanolamine acyltransferase/acyl carrier protein synthetase	L14681	*Escherichia coli*	388	42
ORF615	0603794-0603531	putative				
ORF616	0604413-0603757	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF617	0604549-0605610	3'(2'),5-diphosphonucleoside 3'(2') phosphohydrolase	U33283	*Oryza sativa*	254	45
ORF618	0606619-0605582	leucine dehydrogenase	X79068	*Thermoactinomyces intermedius*	638	49
ORF619	0606843-0607493	inorganic pyrophosphatase	X57545	*Arabidopsis thaliana*	291	37
ORF620	0609068-0608031	beta-ketoacyl-ACP synthase	L13242	*Ricinus communis*	1069	57
ORF621	0609652-0609296	HI0034 homolog	U82598	*Escherichia coli*	196	36
ORF622	0611860-0610109	putative				
ORF623	0611812-0612927	conserved hypothetical protein	AE000579	*Helicobacter pylori*	780	41
ORF624	0613597-0612938	trna delta(2)-isopentenylpyrophosphate transferase	Z98209	*Mycobacterium tuberculosis*	244	37
ORF625	0613895-0613692	delta2-isopentenylpyrophosphate tRNA transferase	Z11831	*Escherichia coli*	134	54
ORF626	0614315-0615244	putative				
ORF627	0615405-0615683	unknown	Z74024	*Mycobacterium tuberculosis*	93	47
ORF628	0617711-0615864	D-alanine:D-alanine ligase	U39788	*Enterococcus hirae*	555	38
ORF629	0618313-0617510	UDP-N-acetylmuramate-alanine ligase (murC)	U32794	*Haemophilus influenzae*	448	47
ORF630	0619335-0618361	transferase, peptidoglycan synthesis (murG)	U32793	*Haemophilus influenzae*	380	39
ORF631	0620416-0619247	spoVE gene product (AA 1-366)	X51419	*Bacillus subtilis*	538	37
ORF632	0619863-0620261	putative				
ORF633	0621184-0620420	hypothetical protein	Y14079	*Bacillus subtilis*	313	44
ORF634	0621690-0621154	murD gene product (AA 1-438)	X51584	*Escherichia coli*	221	43
ORF635	0622399-0621674	MurD	Z95388	*Mycobacterium tuberculosis*	228	41

ORF	Position	Homology	ID	*Species*	Score	I%
ORF636	0623466-0622414	ORF-Y (AA 1-360)	X51584	*Escherichia coli*	543	45
ORF637	0624178-0623570	PROBABLE UDP-N-ACETYLMURAMOYLALANYL-D-GLUTAMYL-2, 6-DIAMINOLIGASE (EC 6.3.2.15).	AB001488	*Bacillus subtilis*	103	43
ORF638	0624918-0624073	UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diamino-pimelate--D-alanyl-D-alanine ligase	X62437	*Synechocystis sp.*	243	33
ORF639	0625346-0626665	chaperonin 60	U56021	*Thermoanaerobacter brockii*	136	31
ORF640	0626514-0626900	putative				
ORF641	0626954-0627853	putative				
ORF642	0627822-0628124	putative				
ORF643	0628715-0628146	elongation factor P	U14003	*Escherichia coli*	467	55
ORF644	0628932-0629801	AMP nucleosidase (EC 3.2.2.4).	D90837	*Escherichia coli*	278	47
ORF645	0630406-0629804	transketolase	Z73234	*Bacillus subtilis*	361	46
ORF646	0630960-0630298	transketolase	Z73234	*Bacillus subtilis*	460	47
ORF647	0631775-0630915	transketolase 1 (TK 1) (tktA)	U32783	*Haemophilus influenzae*	756	47
ORF648	0637635-0638084	alanyl-tRNA synthetase	X59956	*Rhizobium leguminosarum*	436	56
ORF649	0638036-0640207	alanyl-tRNA synthetase	X95571	*Thiobacillus ferrooxidans*	1121	39
ORF650	0640221-0643472	transcription-repair coupling factor (trcF) (mfd)	U32805	*Haemophilus influenzae*	1426	46
ORF651	0640627-0640220	putative				
ORF652	0643485-0644495	uroporphyrinogen decarboxylase	M97208	*Bacillus subtilis*	416	40

ORF	Position	Homology	ID	*Species*	Score	I%
ORF653	0644471-0645430	putative oxygen-independent coproporphyrinogen III oxidase	U06779	*Salmonella typhimurium*	638	43
ORF654	0645394-0645840	oxygen independent coprophorphyrinogen III oxidase	D90912	*Synechocystis sp.*	283	42
ORF655	0645840-0647111	hemY	M97208	*Bacillus subtilis*	133	38
ORF656	0649676-0647109	phosphoprotein	L25078	*Chlamydia trachomatis*	2043	99
ORF657	0649970-0650344	Hc1	M60902	*Chlamydia trachomatis*	603	100
ORF658	0650418-0651722	pCTHom1 gene product	M94254	*Chlamydia trachomatis*	1735	100
ORF659	0651686-0652171	putative				
ORF660	0652516-0652908	phenolhydroxylase component	U32702	*Haemophilus influenzae*	263	41
ORF661	0652799-0653593	phenolhydroxylase component	U32702	*Haemophilus influenzae*	456	51
ORF662	0660136-0661851	YtpT	AF008220	*Bacillus subtilis*	709	52
ORF663	0661740-0662282	spoIIIEB protein	M17445	*Bacillus subtilis*	330	43
ORF664	0662286-0663074	yycJ	D78193	*Bacillus subtilis*	405	38
ORF665	0662951-0663730	C41G7.4	Z81048	*Caenorhabditis elegans*	200	36
ORF666	0664212-0663745	hypothetical protein MTCY180.08	Z97193	*Mycobacterium tuberculosis*	194	38
ORF667	0665619-0664255	D-alanine glycine permease (dagA)	AE000603	*Helicobacter pylori*	205	34
ORF668	0666083-0665727	putative				
ORF669	0666423-0665782	putative				
ORF670	0666831-0668117	putative				
ORF671	0668121-0668375	putative				
ORF672	0668470-0668174	riboflavin synthase beta chain (ribE)	U32810	*Haemophilus influenzae*	192	40
ORF673	0669533-0668616	GTP cyclohydrolase II / 3,4-dihydroxy-2-butanone-4-phoshate synthase	AJ000053	*Arabidopsis thaliana*	800	51

ORF	Position	Homology	ID	*Species*	Score	I%
ORF674	0669868-0669485	unnamed protein product	A38767	*Saccharomyces cerevisiae*	288	49
ORF675	0670780-0669998	ribG gene product	L09228	*Bacillus subtilis*	191	42
ORF676	0671241-0670732	riboflavin-specific deaminase	U27202	*Actinobacillus pleuropneumoniae*	314	51
ORF677	0671182-0672447	seryl-tRNA synthetase	X91007	*Haloarcula marismortui*	736	49
ORF678	0672692-0673231	putative				
ORF679	0673204-0674562	ATPase	L28104	*Transposon Tn5422*	565	41
ORF680	0674612-0675232	unknown	Z74025	*Mycobacterium tuberculosis*	340	43
ORF681	0675327-0676463	rod-shape-determining protein	M22857	*Escherichia coli*	442	37
ORF682	0677003-0676476	biotin [acetyl-CoA carboxylase] ligase	L02354	*Paracoccus denitrificans*	169	49
ORF683	0678422-0677700	ORFX13	L09228	*Bacillus subtilis*	426	43
ORF684	0678717-0679508	2,3-bisphosphoglycerate	M23068	*Homo sapiens*	494	47
ORF685	0679357-0680502	synthesis of [Fe-S] cluster (nifS)	AE000542	*Helicobacter pylori*	150	33
ORF686	0680579-0681280	NifU	AF001780	*Cyanothece PCC 8801*	101	31
ORF687	0681539-0682558	putative				
ORF688	0682554-0683087	putative				
ORF689	0683164-0684465	ORF 4	M72718	*Bacillus subtilis*	708	36
ORF690	0684774-0684418	putative				
ORF691	0684857-0686203	AgX-1 antigen [human, infertile patient, testis, Peptide, 505 aa]	S73498	*Homo sapiens*	338	37
ORF692	0686197-0687204	L-glycerol 3-phosphate dehydrogenase	U00039	*Escherichia coli*	577	38
ORF693	0687341-0688360	putative				
ORF694	0688432-0688193	putative				
ORF695	0689616-0688432	putative				
ORF696	0689960-0689631	putative				
ORF697	0690487-0689846	putative				

ORF	Position	Homology	ID	Species	Score	I%
ORF698	0690717-0690463	putative				
ORF699	0691871-0690672	putative				
ORF700	0693837-0692041	phosphoenolpyruvate carboxykinase	M59372	*Neocallimastix frontalis*	1818	59
ORF701	0694853-0693837	MreB protein	M96343	*Bacillus subtilis*	961	56
ORF702	0697263-0694942	SNF	X98455	*Bacillus cereus*	1073	50
ORF703	0698084-0697170	putative				
ORF704	0698392-0697979	putative				
ORF705	0698822-0700117	trigger factor (tig)	AE000591	*Helicobacter pylori*	84	34
ORF706	0700287-0700895	proteosome major subunit	AF013216	*Myxococcus xanthus*	615	59
ORF707	0700912-0702165	ATP-dependent protease ATPase subunit	L18867	*Escherichia coli*	1183	55
ORF708	0702183-0703412	poly(A) polymerase	L47709	*Bacillus subtilis*	362	38
ORF709	0703522-0705000	hypothetical protein	D90912	*Synechocystis sp.*	809	41
ORF710	0705011-0705604	putative				
ORF711	0706159-0705704	Preprotein translocase subunit	AF022186	*Cyanidium caldarium*	165	44
ORF712	0706521-0706138	secA	X99401	*Bacillus firmus*	155	42
ORF713	0708103-0706496	SecA	U66081	*Mycobacterium smegmatis*	1044	58
ORF714	0708398-0708078	cp-SecA; chloroplast SecA homolog	U71123	*Zea mays*	258	69
ORF715	0708607-0708248	SecA	U21192	*Streptomyces lividans*	179	42
ORF716	0710278-0708872	putative				
ORF717	0711164-0710262	phosphatidylserine decarboxylase	U72715	*Chlamydia trachomatis*	1548	99
ORF718	0711432-0712763	homologous to E.coli 50K	X62539	*Bacillus subtilis*	713	54
ORF719	0712767-0713438	ultraviolet N-glycosylase/AP lyase	U22181	*Micrococcus luteus*	273	45
ORF720	0714232-0713651	putative				
ORF721	0714632-0714120	putative				
ORF722	0715592-0714834	putative				
ORF723	0715854-0715558	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF724	0716937-0715921	putative				
ORF725	0718357-0717149	3-phosphoglycerate kinase	U83197	*Chlamydia trachomatis*	2049	100
ORF726	0718500-0718862	putative				
ORF727	0719797-0718499	phosphate permease (YBR296C)	U32834	*Haemophilus influenzae*	997	42
ORF728	0720273-0719782	putative				
ORF729	0720452-0720144	H. influenzae predicted coding region HI1603	U32834	*Haemophilus influenzae*	164	37
ORF730	0720613-0721575	dciAD	X56678	*Bacillus subtilis*	722	41
ORF731	0721559-0722356	was dppE	U00039	*Escherichia coli*	477	44
ORF732	0723248-0722397	chromosome partitioning protein ParB	U87804	*Caulobacter crescentus*	388	50
ORF733	0724598-0723378	NifS protein.	D90811	*Escherichia coli*	805	39
ORF734	0725763-0724576	hypothetical protein	D64004	*Synechocystis sp.*	154	41
ORF735	0726516-0725767	Multidrug resistance protein 1 (P-glycoprotein 1).	D90811	*Escherichia coli*	607	54
ORF736	0726819-0726538	ABC transporter subunit	D64004	*Synechocystis sp.*	266	58
ORF737	0727493-0726753	ABC transporter subunit	D64004	*Synechocystis sp.*	854	71
ORF738	0727984-0727469	ABC transporter subunit	D64004	*Synechocystis sp.*	531	55
ORF739	0728778-0728329	putative				
ORF740	0729346-0728759	antiviral protein	L36940	*Saccharomyces cerevisiae*	115	33
ORF741	0732426-0729442	penicillin-binding protein 2 (pbp2)	U32688	*Haemophilus influenzae*	208	43
ORF742	0733246-0734427	major outer membrane protein precursor	M14738	*Chlamydia trachomatis*	2045	99
ORF743	0734814-0735659	ribosomal protein S2	U60196	*Chlamydia trachomatis*	1269	76
ORF744	0735659-0736504	elongation factor Ts	U60196	*Chlamydia trachomatis*	1278	90
ORF745	0736520-0737254	UMP kinase	U60196	*Chlamydia trachomatis*	1153	94
ORF746	0737254-0737787	ribosome-releasing factor	U60196	*Chlamydia trachomatis*	760	92
ORF747	0737942-0738679	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF748	0738838-0739740	ORF3; putative 39 kDa protein	U40604	*Listeria monocytogenes*	116	31
ORF749	0742057-0740060	XcpQ	X68594	*Pseudomonas aeruginosa*	453	37
ORF750	0742869-0742045	putative				
ORF751	0743378-0742824	putative				
ORF752	0744298-0743306	unknown	Z80233	*Mycobacterium tuberculosis*	137	40
ORF753	0744714-0744430	putative	M69228	*Caulobacter crescentus*	117	38
ORF754	0744985-0744611	putative				
ORF755	0745557-0744958	putative				
ORF756	0746412-0745561	putative				
ORF757	0746772-0746416	putative				
ORF758	0748269-0746944	PscN	AF010151	*Pseudomonas aeruginosa*	1220	55
ORF759	0748966-0748274	putative				
ORF760	0749426-0748965	putative				
ORF761	0749702-0749433	putative				
ORF762	0750029-0749721	putative				
ORF763	0752307-0750007	putative				
ORF764	0752913-0752503	putative				
ORF765	0754587-0753616	NAD(P)H:glutamyl-transfer RNA reductase	M57676	*Bacillus subtilis*	172	40
ORF766	0755000-0756814	DNA gyrase subunit B	U35453	*Clostridium acetobutylicum*	970	38
ORF767	0756796-0758301	gyrA	X92503	*Mycobacterium smegmatis*	409	49
ORF768	0758691-0758446	unknown	Z74024	*Mycobacterium tuberculosis*	107	34
ORF769	0759775-0759338	SfhB	U50134	*Escherichia coli*	241	48
ORF770	0760242-0759871	putative				
ORF771	0760538-0760188	putative				
ORF772	0760966-0761772	3-deoxy-D-manno-octulosonate 8-phosphate synthetase	U72493	*Chlamydia trachomatis*	1350	99

ORF	Position	Homology	ID	*Species*	Score	I%
ORF773	0761759-0762142	unknown	U72493	*Chlamydia trachomatis*	536	94
ORF774	0762267-0762983	ATP binding protein	U72493	*Chlamydia trachomatis*	1197	99
ORF775	0764465-0763335	chlanectin coding region	M17875	*Chlamydia trachomatis*	239	100
ORF776	0764857-0764438	putative				
ORF777	0766068-0764821	unknown function	Z32530	*Chlamydia trachomatis*	1803	99
ORF778	0766475-0766065	unknown function	Z32530	*Chlamydia trachomatis*	704	100
ORF779	0767989-0766934	RecA	U16739	*Chlamydia trachomatis*	1753	100
ORF780	0768785-0768252	unknown function	Z32530	*Chlamydia trachomatis*	904	99
ORF781	0770062-0768791	unknown function	Z32530	*Chlamydia trachomatis*	2249	100
ORF782	0770138-0770470	putative				
ORF783	0770661-0770185	putative				
ORF784	0770924-0770634	putative				
ORF785	0772010-0771330	putative				
ORF786	0772390-0773391	unknown	D26185	*Bacillus subtilis*	486	35
ORF787	0774221-0773427	ORF_f169	U18997	*Escherichia coli*	263	51
ORF788	0775996-0774191	DNA topoisomerase I	L27797	*Bacillus subtilis*	1357	52
ORF789	0776663-0777706	putative				
ORF790	0777195-0776953	putative				
ORF791	0779222-0777732	ORF_f397	U29581	*Escherichia coli*	93	40
ORF792	0779321-0781552	putative				
ORF793	0781297-0782442	putative				
ORF794	0782447-0785524	exonuclease V (AA 1-1180)	X04581	*Escherichia coli*	557	49
ORF795	0785532-0786002	putative				
ORF796	0786412-0785546	MreC protein	M31792	*Escherichia coli*	81	64
ORF797	0787741-0786611	aspartate aminotransferase precursor	M12105	*Gallus gallus*	700	42
ORF798	0787620-0788021	putative				
ORF799	0790124-0787920	GreA	U02878	*Rickettsia prowazekii*	84	33
ORF800	0790160-0790609	putative				
ORF801	0790634-0792016	NADH:ubiquinone oxidoreductase subunit A	Z37111	*Vibrio alginolyticus*	409	37
ORF802	0793084-0792059	delta_aminolevulinic acid dehydratase	L24386	*Bradyrhizobium japonicum*	867	52

ORF	Position	Homology	ID	*Species*	Score	I%
ORF803	0793343-0794056	putative				
ORF804	0794046-0794957	putative				
ORF805	0795401-0795144	putative				
ORF806	0795575-0796255	ompR gene product	X92405	*Neisseria meningitidis*	103	32
ORF807	0796278-0797015	glucose-1-phosphate thymidylyltransferase	U67553	*Methanococcus jannaschii*	216	36
ORF808	0796985-0797365	YqiD	D84432	*Bacillus subtilis*	184	58
ORF809	0797260-0797856	farnesyl diphosphate synthase	D13293	*Bacillus stearothermophilus*	107	37
ORF810	0797772-0798086	putative				
ORF811	0798426-0797935	Orf39.9	X61000	*Escherichia coli*	290	51
ORF812	0798925-0798416	This ORF is homologous to a 40.0 kd hypothetical protein in the htrB 3' region from E. coli, Accession Number X61000	L22217	*Mycoplasma-like organism*	150	46
ORF813	0799301-0799927	ribosomal protein S4 (rps4)	AE000633	*Helicobacter pylori*	407	46
ORF814	0800862-0800029	apurinic/apyrimidinic endonuclease	U40707	*Caenorhabditis elegans*	397	35
ORF815	0801065-0802129	mviB homolog	U50732	*Chlamydia trachomatis*	1716	97
ORF816	0802023-0802673	mviB homolog	U50732	*Chlamydia trachomatis*	973	97
ORF817	0802920-0803246	lorf2; possible membrane-bound protein	U50732	*Chlamydia trachomatis*	280	100
ORF818	0803105-0804220	76 kDa protein	L23921	*Chlamydia pneumoniae*	775	59
ORF819	0804307-0805356	putative				
ORF820	0805290-0806282	76 kDa protein	L23921	*Chlamydia pneumoniae*	125	50
ORF821	0806453-0808081	putative				
ORF822	0808026-0809009	putative				
ORF823	0810461-0809079	putative				
ORF824	0811605-0810328	putative				
ORF825	0811725-0812342	putative				
ORF826	0812329-0813522	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF827	0813455-0813772	putative				
ORF828	0813732-0814334	putative				
ORF829	0815213-0814314	putative				
ORF830	0814878-0814396	putative				
ORF831	0815688-0815428	30S ribosomal protein S20	Z67753	*Odontella sinensis*	150	38
ORF832	0816116-0817456	KIAA0336	AB002334	*Homo sapiens*	90	32
ORF833	0817608-0819320	RNA polymerase sigma-subunit	J05546	*Chlamydia trachomatis*	2868	100
ORF834	0819324-0819713	putative				
ORF835	0819704-0820402	dihydropterin pyrophosphokinase /dihydropteroate synthase	Y08611	*Pisum sativum*	310	45
ORF836	0820375-0821061	dihydropteroate synthase	X68068	*Neisseria meningitidis*	100	48
ORF837	0821061-0821537	dihydrofolate reductase	Z84379	*Streptococcus pneumoniae*	168	45
ORF838	0821646-0822239	M. jannaschii predicted coding region MJ0768	U67522	*Methanococcus jannaschii*	139	41
ORF839	0822182-0822931	putative				
ORF840	0824355-0823045	nitrogen metabolism regulator	M58480	*Thiobacillus ferrooxidans*	133	58
ORF841	0825894-0824359	helicase	M63176	*Staphylococcus aureus*	893	50
ORF842	0826259-0825879	helicase	M63176	*Staphylococcus aureus*	282	47
ORF843	0826340-0827026	ipa-57d gene product	X73124	*Bacillus subtilis*	602	52
ORF844	0827014-0827250	putative				
ORF845	0827856-0827230	hypothetical	U32712	*Haemophilus influenzae*	302	45
ORF846	0828007-0829275	19/20 residue stretch (32-51) identical to N-terminal putative signal sequence of unknown, partly cloned B. subtilis gene.; putative	L19954	*Bacillus subtilis*	442	37
ORF847	0829355-0830953	heat shock protein GroEL	U55047	*Bradyrhizobium japonicum*	418	36

ORF	Position	Homology	ID	*Species*	Score	I%
ORF848	0831119-0831748	basl protein	Z34917	*Hordeum vulgare*	516	47
ORF849	0832152-0831751	putative				
ORF850	0832744-0832214	putative				
ORF851	0833446-0832805	putative				
ORF852	0833802-0833368	putative				
ORF853	0834679-0833879	putative				
ORF854	0835452-0834661	putative				
ORF855	0835778-0835371	putative				
ORF856	0836482-0835775	putative				
ORF857	0836602-0837264	putative				
ORF858	0837209-0838699	putative				
ORF859	0838760-0839575	putative				
ORF860	0839942-0840583	putative				
ORF861	0840445-0841713	putative				
ORF862	0841659-0842459	putative				
ORF863	0842523-0843068	putative				
ORF864	0843495-0843031	putative				
ORF865	0843239-0846196	putative				
ORF866	0844137-0843802	putative				
ORF867	0848043-0846217	putative				
ORF868	0850123-0848150	putative				
ORF869	0851645-0850230	putative				
ORF870	0853696-0851669	putative				
ORF871	0854836-0853700	putative				
ORF872	0855525-0854920	putative				
ORF873	0856240-0855437	putative				
ORF874	0857183-0856233	putative				
ORF875	0859439-0857451	putative				
ORF876	0859946-0859587	putative				
ORF877	0859642-0860640	putative				
ORF878	0861599-0860724	putative				
ORF879	0862053-0861580	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF880	0863540-0862098	putative				
ORF881	0863930-0863571	putative				
ORF882	0864697-0863996	putative				
ORF883	0864938-0866248	DNA mismatch repair protein (mutL)	U32692	*Haemophilus influenzae*	506	47
ORF884	0866303-0866605	putative				
ORF885	0866665-0867732	YqhT	D84432	*Bacillus subtilis*	444	39
ORF886	0867810-0869090	putative				
ORF887	0869094-0869357	putative				
ORF888	0869270-0871372	fimbrial assembly protein	L13865	*Pseudomonas aeruginosa*	181	40
ORF889	0871299-0872582	xpsE gene product	X59079	*Xanthomonas campestris*	825	56
ORF890	0872429-0872860	secretion protein XcpR	Y09102	*Acinetobacter calcoaceticus*	213	48
ORF891	0872875-0873915	ORF_o398	U18997	*Escherichia coli*	271	33
ORF892	0873812-0873360	putative				
ORF893	0874028-0874438	putative				
ORF894	0874778-0875386	putative				
ORF895	0875774-0876382	putative				
ORF896	0877809-0877000	secretion system apparatus, SsaT	X99944	*Salmonella typhimurium*	174	34
ORF897	0878151-0877876	yscS	L25667	*Yersinia pseudotuberculosis*	172	44
ORF898	0878846-0878172	pathogenicity protein	M64094	*Xanthomonas campestris*	464	46
ORF899	0878883-0879161	putative				
ORF900	0879773-0879105	PscL	U56077	*Pseudomonas aeruginosa*	141	34
ORF901	0880885-0880052	putative				
ORF902	0881830-0880889	HrcJ	U56662	*Erwinia amylovora*	236	43
ORF903	0882904-0881948	ORF YOR196c	Z75104	*Saccharomyces cerevisiae*	685	44
ORF904	0883794-0882901	dihydrolipoamide dehydrogenase	L31844	*Clostridium magnum*	578	38
ORF905	0884296-0883661	YqiV	D84432	*Bacillus subtilis*	437	44
ORF906	0884996-0884508	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF907	0888777-0885166	helicase of the snf2/rad54 family	D90916	*Synechocystis sp.*	824	43
ORF908	0890148-0888940	sodium-coupled branched-chain amino acid carrier	D49784	*Clostridium perfringens*	230	35
ORF909	0891164-0890325	putative Fmu protein	Y13937	*Bacillus subtilis*	220	41
ORF910	0891463-0891116	putative				
ORF911	0893278-0891968	DD-carboxypeptidase	M85047	*Bacillus subtilis*	302	39
ORF912	0893356-0893808	putative				
ORF913	0893909-0893643	putative				
ORF914	0894258-0893821	hypothetical protein	D90908	*Synechocystis sp.*	155	39
ORF915	0894778-0894248	putative				
ORF916	0895892-0895050	putative				
ORF917	0895951-0896829	putative				
ORF918	0900783-0897064	DNA polymerase III alpha-subunit (dnaE)	AE000646	*Helicobacter pylori*	1974	43
ORF919	0902032-0900791	UhpC protein	M17102	*Escherichia coli*	1117	52
ORF920	0902677-0903876	histidine--tRNA ligase	Z17214	*Streptococcus equisimilis*	686	47
ORF921	0903731-0903471	putative				
ORF922	0903860-0905605	aspartyl-tRNA synthetase	D90910	*Synechocystis sp.*	1339	51
ORF923	0905746-0906474	mip-like protein	X66126	*Chlamydia trachomatis*	1196	98
ORF924	0906589-0906945	spoU	L40369	*Chlamydia trachomatis*	607	100
ORF925	0907306-0907001	trxA	L39892	*Chlamydia psittaci*	380	76
ORF926	0908101-0908742	putative				
ORF927	0908721-0909194	hypothetical protein	D90914	*Synechocystis sp.*	150	37
ORF928	0909198-0909584	DNA polymerase III	Z48003	*Staphylococcus aureus*	181	40
ORF929	0909583-0909951	putative				
ORF930	0910081-0910569	VdlD	U94318	*Helicobacter pylori*	197	43
ORF931	0910615-0910944	putative				
ORF932	0910948-0912261	acid-inducible gene	L13845	*Sinorhizobium meliloti*	145	50
ORF933	0912399-0912629	putative				
ORF934	0912595-0913218	UDP-3-O-acyl-GlcNAc deacetylase	U67855	*Pseudomonas aeruginosa*	309	39

ORF	Position	Homology	ID	*Species*	Score	I%
ORF935	0913203-0913676	(3R)-hydroxymyristol acyl carrier protein dehydrase	D90910	*Synechocystis sp.*	302	59
ORF936	0913691-0914485	UDP-N-acetylglucosamine acyltransferase	L22690	*Rickettsia rickettsii*	503	38
ORF937	0914516-0915136	methionyl-tRNA formyltransferase	X63666	*Escherichia coli*	407	42
ORF938	0915144-0915467	putative				
ORF939	0915629-0916633	putative				
ORF940	0916051-0916539	putative				
ORF941	0916965-0917627	ribosomal protein L3 (rpL3)	U32761	*Haemophilus influenzae*	470	48
ORF942	0917639-0918304	50S ribosomal protein L4	AB000111	*Synechococcus sp.*	210	43
ORF943	0918377-0918655	ribosomal protein L23	Z21677	*Thermotoga maritima*	116	47
ORF944	0918682-0919533	rpl2	M74770	*Mycoplasma-like organism*	800	48
ORF945	0919542-0919829	Mycoplasma pneumoniae, ribosomal protein S19; similar to GenBank Accession Number S36895, from M. bovis	AE000061	*Mycoplasma pneumoniae*	315	68
ORF946	0919738-0920157	ribosomal protein L22	Z21677	*Thermotoga maritima*	240	49
ORF947	0920184-0920840	ribosomal protein S3 (rpS3)	U32761	*Haemophilus influenzae*	605	57
ORF948	0920866-0921294	ribosomal protein L16	Z21677	*Thermotoga maritima*	434	62
ORF949	0921299-0921514	ribosomal protein CtrL29e	M80325	*Chlamydia trachomatis*	343	99
ORF950	0921510-0921758	ribosomal protein S17e	M80325	*Chlamydia trachomatis*	419	100
ORF951	0921778-0922143	ribosomal protein CtrL14e	M80325	*Chlamydia trachomatis*	618	100
ORF952	0922159-0922491	ribosomal protein CtrL24e	M80325	*Chlamydia trachomatis*	568	100
ORF953	0922571-0923035	ribosomal protein CtrL5e	M80325	*Chlamydia trachomatis*	793	99
ORF954	0923160-0923453	ribosomal protein CtrS8e	M80325	*Chlamydia trachomatis*	487	98
ORF955	0923484-0924032	ribosomal protein L6	M60652	*Chlamydia trachomatis*	927	100
ORF956	0924057-0924425	ribosomal protein CtrL18e	M80325	*Chlamydia trachomatis*	605	99
ORF957	0924443-0924937	ribosomal protein CtrS5e	M80325	*Chlamydia trachomatis*	814	99

ORF	Position	Homology	ID	*Species*	Score	I%
ORF958	0924933-0925364	ribosomal protein CtrL15e	M80325	*Chlamydia trachomatis*	740	99
ORF959	0925390-0926760	homolog	L25077	*Chlamydia trachomatis*	2254	99
ORF960	0926819-0927184	ribosomal protein S13	L33834	*Chlamydia trachomatis*	604	100
ORF961	0927209-0927604	ribosomal protein S11	L33834	*Chlamydia trachomatis*	646	98
ORF962	0927628-0928155	RNA polymerase alpha-subunit	L33834	*Chlamydia trachomatis*	847	97
ORF963	0928100-0928759	RNA polymerase alpha-subunit	L33834	*Chlamydia trachomatis*	1040	98
ORF964	0929222-0930244	glyceraldehyde-3-phosphate dehydrogenase	U83198	*Chlamydia trachomatis*	1735	99
ORF965	0930222-0930656	putative				
ORF966	0930608-0931078	putative				
ORF967	0931367-0931666	putative				
ORF968	0931549-0931959	putative				
ORF969	0932070-0932579	crossover junction endodeoxyribonuclease (ruvC)	U32717	*Haemophilus influenzae*	250	41
ORF970	0932602-0933201	Holliday junction DNA helicase (ruvA)	U32716	*Haemophilus influenzae*	258	38
ORF971	0933364-0933621	nucleoside diphosphate kinase (ndk)	AE000540	*Helicobacter pylori*	264	60
ORF972	0933522-0933785	nucleoside 5'-diphosphate phosphotransferase (EC 2.7.4.6)	J05207	*Myxococcus xanthus*	186	64
ORF973	0934546-0933848	hypothetical protein (GB:U14003_297)	U39706	*Mycoplasma genitalium*	156	36
ORF974	0936368-0934539	homologous to E.coli gidA	X62540	*Pseudomonas putida*	1562	51
ORF975	0938063-0936666	replicative DNA helicase	D26185	*Bacillus subtilis*	848	41
ORF976	0938538-0939098	phosphatidylglycerophosphate synthase (pgsA)	AE000610	*Helicobacter pylori*	120	33
ORF977	0939329-0940933	adenine nucleotide translocase	Z49227	*Arabidopsis thaliana*	668	40
ORF978	0941076-0942068	putative protease	AF008220	*Bacillus subtilis*	265	36

ORF	Position	Homology	ID	*Species*	Score	I%
ORF979	0942088-0944685	DNA polymerase	D12982	*Bacillus caldotenax*	1334	42
ORF980	0944634-0945287	T05G5.5	Z27079	*Caenorhabditis elegans*	198	32
ORF981	0945434-0946294	'The first ATG in the open reading frame was chosen as the initiation codon.'	L27278	*Pseudomonas fluorescens*	882	68
ORF982	0946293-0946676	'The first GTG in the open reading frame was chosen as the initiation codon.'	L27276	*Deinococcus radiodurans*	417	65
ORF983	0947105-0948454	ADPglucose pyrophosphorylase	M31616	*Oryza sativa*	755	44
ORF984	0948522-0949277	putative				
ORF985	0949277-0949594	YlbH protein	Z98682	*Bacillus subtilis*	223	41
ORF986	0949849-0950676	putative				
ORF987	0950680-0951330	ferrochelatase	M59288	*Mus musculus*	260	42
ORF988	0951281-0951643	ferrochelatase	D26106	*Cucumis sativus*	178	47
ORF989	0951788-0952798	putative				
ORF990	0953581-0954264	putative				
ORF991	0954426-0955157	putative				
ORF992	0955754-0957940	orf4 gene product	X93084	*Methanosarcina barkeri*	130	41
ORF993	0957837-0959312	OppB gene product	X56347	*Bacillus subtilis*	327	38
ORF994	0959299-0961050	dipeptide ABC transporter, permease protein (dppC)	AE000548	*Helicobacter pylori*	263	39
ORF995	0961514-0961053	methylated DNA protein cysteine methyltransferase	U67593	*Methanococcus jannaschii*	109	39
ORF996	0962575-0961487	putative				
ORF997	0961979-0961584	putative				
ORF998	0964914-0962545	phenylalanyl-tRNA synthetase beta subunit	Z75208	*Bacillus subtilis*	775	37
ORF999	0964941-0965708	putative				
ORF1000	0967023-0966193	unknown	Z48008	*Saccharomyces cerevisiae*	492	44
ORF1001	0967444-0968061	putative				
ORF1002	0968903-0968064	putative				

ORF	Position	Homology	ID	*Species*	Score	I%
ORF1003	0970685-0969528	transcriptional activator of pilA	Z12154	*Pseudomonas aeruginosa*	849	45
ORF1004	0971806-0971024	sensor protein	L39904	*Myxococcus xanthus*	147	30
ORF1005	0973053-0972388	putative				
ORF1006	0974546-0973746	unknown	D64126	*Bacillus subtilis*	500	50
ORF1007	0975223-0974558	unknown	D26185	*Bacillus subtilis*	141	44
ORF1008	0975989-0975207	hypothetical protein in htrA-dapD intergenic region	AE000126	*Escherichia coli*	142	42
ORF1009	0976520-0976254	unknown	Z49939	*Saccharomyces cerevisiae*	183	39
ORF1010	0976588-0976899	putative				
ORF1011	0976886-0977635	peptide release factor 2	X99401	*Bacillus firmus*	534	44
ORF1012	0977661-0977933	release factor 2	AF013188	*Bacillus subtilis*	187	52
ORF1013	0977918-0978433	putative				
ORF1014	0978619-0978984	spore coat protein CotRC	D50551	*Bacillus subtilis*	355	52
ORF1015	0978933-0979331	hypothetical	U32717	*Haemophilus influenzae*	199	40
ORF1016	0981197-0979389	putative				
ORF1017	0979711-0980112	putative				
ORF1018	0982116-0981148	putative				
ORF1019	0982321-0983598	UDP-N-acetylglucosamine enolpyruvyl transferase (murZ)	U32788	*Haemophilus influenzae*	593	38
ORF1020	0984488-0983862	arginyl-tRNA-synthetase	D64006	*Synechocystis sp.*	347	44
ORF1021	0985381-0984371	arginyl-tRNA-synthetase	D64006	*Synechocystis sp.*	782	58
ORF1022	0986103-0985399	hypothetical protein	D90915	*Synechocystis sp.*	224	35
ORF1023	0986693-0986046	No definition line found	U00021	*Mycobacterium leprae*	286	50
ORF1024	0987607-0986693	o298; This 298 aa ORF is 33 pct identical (24 gaps) to 248 residues of an approx. 256 aa protein CDSA_ECOLI SW: P06466	AE000238	*Escherichia coli*	132	46

ORF	Position	Homology	ID	*Species*	Score	I%
ORF1025	0988119-0987616	conserved hypothetical protein	AE000627	*Helicobacter pylori*	343	49
ORF1026	0988253-0987936	hypothetical protein (HI0920)	U67577	*Methanococcus jannaschii*	110	38
ORF1027	0988831-0989163	putative				
ORF1028	0989693-0993442	protein-export membrane protein SecD	D64000	*Synechocystis sp.*	447	38
ORF1029	0993408-0993785	protein-export membrane protein	U83136	*Rhodobacter sphaeroides*	240	43
ORF1030	0993835-0993416	putative				
ORF1031	0993882-0994262	putative				
ORF1032	0994226-0995656	RecJ recombination protein	U41759	*Chlamydia psittaci*	880	66
ORF1033	0996063-0996611	unknown	U41759	*Chlamydia psittaci*	533	75
ORF1034	0996885-0998267	glutamyl-tRNA synthetase homolog	U41759	*Chlamydia psittaci*	2018	83
ORF1035	0998962-0999225	9-kDa cysteine-rich outer membrane protein	M35148	*Chlamydia trachomatis*	504	100
ORF1036	0999393-1001033	outer membrane protein 2	M23001	*Chlamydia trachomatis*	2857	100
ORF1037	1001214-1001516	15-kDa serine-rich outer membrane protein	M35148	*Chlamydia trachomatis*	276	94
ORF1038	1001392-1001664	15-kDa serine-rich outer membrane protein	M35148	*Chlamydia trachomatis*	438	97
ORF1039	1003697-1001823	ORF of prc gene (alt.)	D00674	*Escherichia coli*	486	42
ORF1040	1004477-1004845	StrA	M86701	*Haemophilus influenzae*	454	70
ORF1041	1004990-1005382	ribosomal protein S7	Z11567	*Chlamydia trachomatis*	662	99
ORF1042	1005415-1007496	translation elongation factor EF-G (fusA)	AE000625	*Helicobacter pylori*	2147	62
ORF1043	1007507-1007821	ribosomal protein S10	Z21676	*Spirulina platensis*	350	68
ORF1044	1007802-1008698	NADPH-sulfite reducatase flavoprotein component	M23008	*Escherichia coli*	113	48
ORF1045	1009381-1009121	unknown	Z92774	*Mycobacterium tuberculosis*	102	42

ORF	Position	Homology	ID	*Species*	Score	I%
ORF1046	1010648-1012054	serine hydroxymethyltransferase	Z38002	*Bacillus subtilis*	1021	55
ORF1047	1012397-1011942	putative				
ORF1048	1012042-1012635	ATP-dependent Clp protease proteolytic subunit	D90915	*Synechocystis sp.*	365	44
ORF1049	1012593-1012862	putative				
ORF1050	1012811-1013440	diaminopimelate epimerase (dapF)	U32759	*Haemophilus influenzae*	108	40
ORF1051	1013456-1014055	putative				
ORF1052	1013977-1014489	putative				
ORF1053	1015224-1014529	hypothetical 28.1 kD protein in udp-rfaH intergenic region	AE000459	*Escherichia coli*	263	38
ORF1054	1016002-1015145	putative				
ORF1055	1016994-1015939	conserved hypothetical protein	AE000579	*Helicobacter pylori*	428	42
ORF1056	1017766-1017245	putative				
ORF1057	1018911-1017916	putative				
ORF1058	1019191-1018580	putative				
ORF1059	1020199-1019831	hemolysin	AE000647	*Helicobacter pylori*	164	33
ORF1060	1021007-1020114	unknown	Z95208	*Mycobacterium tuberculosis*	201	36
ORF1061	1021569-1021075	putative				
ORF1062	1022411-1022097	putative				
ORF1063	1023347-1023667	50S ribosomal subunit protein L21	U18997	*Escherichia coli*	218	43
ORF1064	1023701-1023949	50S ribosomal protein L27	U38804	*Porphyra purpurea*	251	64
ORF1065	1024042-1024776	ORF_f390	U18997	*Escherichia coli*	603	51
ORF1066	1024704-1025045	GTP-binding protein (obg)	U32769	*Haemophilus influenzae*	161	37
ORF1067	1025881-1024967	hypothetical protein	D90903	*Synechocystis sp.*	439	35
ORF1068	1026546-1025839	YcdI	AB000617	*Bacillus subtilis*	312	40
ORF1069	1027379-1026546	adhesion protein	D90903	*Synechocystis sp.*	354	35

ORF	Position	Homology	ID	*Species*	Score	I%
ORF1070	1030604-1027929	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	95	49
ORF1071	1033249-1030508	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	75	36
ORF1072	1031733-1032086	putative				
ORF1073	1037037-1033456	putative 98 kDa outer membrane protein	U72499	*Chlamydia psittaci*	160	46
ORF1074	1035674-1035910	putative				
ORF1075	1036175-1036507	putative				
ORF1076	1038592-1036967	putative				

References

Altschul, S.F. et al., 1990, J. Mol. Biol., 215 : 403-410.

Arlinghaus, H.F. et al., 1997, Anal. Biochem., 69, 18, 3747-53.

Barany, F., 1911, Proc. Natl. Acad. Sci. USA, 88 : 189-193.

Beattie, K. et al., 1993, Clin. Chem., 39(4) : 719-721.

Borman, S., 1996, Chem. Eng. News, 74(50) : 42-43.

Buckholz, R.G., 1993, Curr. Op. Biotechnology, 4 : 538-542.

Burg, J.L. et al., 1996, Mol. and Cell. Probes, 10 :257-271.

Casas-Ciria J. et al., 1996

Chatelier, R.C. et al., 1995, Anal. Biochem., 229, 1, 112-118.

Chee, M. et al., 1996, Science, 274 : 610-613.

Chu, B.C.F. et al., 1986, Nucleic Acids Research, 14:5591-5603.

Cserzo, M., Wallin, E., Simon, I. von Heijne G and Elofsson, A., 1997, Prot. Eng., 10 : 673-676.

Derisi, J. et al., 1996, Nature Genet, 14 : 457-460.

Duck, P. et al., 1990, Biotechniques, 9 : 142-147.

Edwards, C.P., and Aruffo, A., 1993, Curr. Op. Biotechnology 4 : 558-563.

Erlich, H.A., 1989, In PCR Technology. Principles and Applications for DNA Amplification. New York : Stockton Press.

Felgner, et al., 1987, Proc. Natl. Acad. Sci. USA, 84 : 7413.

Fodor, S.P.A. et al., 1991, Science, 251 : 767-771.

Fraley et al., 1980, J. Biol. Chem., 255 : 10431.

Guateli, J.C. et al., 1990, Proc. Natl. Acad. Sci. USA, 87 : 1874-1878.

Hackstadt T., Trends Microbiol. 1996 5: 288-293

Houbenweyl, 1974, in Meuthode der Organischen Chemie, E. Wunsch Ed., Volume 15-I et 15-II, Thieme, Stuttgart.

Huygen, K. et al., 1996, Nature Medicine, 2(8) : 893-898.

Innis, M.A. et al. 1990, in PCR Protocols. A guide to Methods and Applications. San Diego : Academic Press.

Kaneda, et al., 1989, Science, 243 : 375.

Kievitis, T. et al., 1991, J. Virol. Methods, 35 : 273-286.

Kohler, G. et al., 1975, Nature, 256(5517) : 495-497.

Krone, J.R. et al., 1997, Anal. Biochem., 244, 1, 124-132.

Kwoh, D.Y. et al., 1989, Proc. Natl. Acad. Sci. USA, 86: 1173-1177.

Lipshutz, R.J. et al., 1995, Biotechniques, 19(3) : 442-447.

Livache, T. et al., 1994, Nucleic Acids Research, 22(15) : 2915-2921.

Lockhart, D.J. et al., 1996, Nature Biotechnol., 14 : 1675-1680.

Luckow, V.A., 1993, Curr. Op. Biotechnology 4 : 564-572.

Matson, R.S. et al., 1994, Anal. Biochem., 217 : 306-310.

Matthews, J.A. et al., 1988, Anal. Biochem., 169 : 1-25.

Mérel, P., 1994, Biofutur, 139 : 58.

Merrifield, R.D., 1966, J. Am. Chem. Soc., 88(21) : 5051-5052.

Midoux, 1993, Nucleic Acids Research, 21 : 871-878.

Miele, E.A. et al., 1983, J. Mol. Biol., 171 : 281-295.

O'Donell-Maloney, M.J., 1996, Trends Biotechnol., 14 : 401-407.

Pagano et al., 1967, J. Virol., 1 : 891.

Raulston JE., Mol Bicrobiol 1995 15:607-616

Rolfs, A. et al., 1991, In PCR Topics. Usage of Polymerase Chain reaction in Genetic and Infectious Disease. Berlin : Springer-Verlag.

Sambrook, J. et al., 1989, In Molecular cloning : A Laboratory Manual. Cold Spring Harbor, NY : Cold Spring Harbor Laboratory Press.

Sanchez-Pescador, R., 1988, J. Clin. Microbiol., 26(10) : 1934-1938.

Segev D., 1992, in "Non-radioactive Labeling and Detection of Biomolecules". Kessler C. Springer Verlag, Berlin, New-York : 197-205.

Sheldon, E.L., 1993, Clin. Chem., 39(4) : 718-719.

Shiver, J.W., 1995, in Vaccines, eds Chanock, R.M. Brown, F. Ginsberg, H.S. & Norrby, E., pp.95-98, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shoemaker, D.D. et al., 1996, Nature Genet., 14 : 450-456.

Sosnowsky et al., 1997, Proc. Natl. Acac. Sci, USA, 94, 1119-1123.

Sundelof, et al., 1993 Scand. J. Infec. Dis., 25 :259-261.

Tascon, R.E. et al., 1996, Nature Medicine, 2(8) : 888-892.

Urdea, M.S., 1988, Nucleic Acids Research, 11 : 4937-4957.

Walker, G.T. et al., 1992, Nucleic Acids Research, 20 : 1691-1696.

Walker, G.T. et al., 1992, Proc. Natl. Acad. Sci. USA, 89 : 392-396.

White, B.A. et al., 1997, Methods in Molecular Biology, 67, Humana Press, Towota.

Yershov, G. et al., 1996, Proc. Natl. Acad. Sci. USA, 93 : 4913-4918.

In the sequence description SEQ ID No. 1 of the sequence listing below, the number which is indicated below the line listing the nucleotides should be taken as the position of the last nucleotide in the said line. For example, the number 60 which is indicated below the first line listing the nucleotides of the sequence SEQ ID No. 1 corresponds to the position of the last nucleotide in the first line.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT
(A) NAME: GENSET
(B) STREET: 24 RUE ROYALE
(C) CITY: PARIS
(D) COUNTRY: FRANCE
(E) POSTAL CODE: 75008

(ii) TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN PARTICULAR FOR THE DIAGNOSIS, PREVENTION AND TREATMENT OF INFECTION

(iii) NUMBER OF SEQUENCES: 1076

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Genlist

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1038608 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(iii) MOLECULE TYPE: genomic DNA

(vi) ORIGINAL SOURCE:
(1.vi.A) ORGANISM: Chlamydia trachomatis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

gaagcaaaac cctgaaactg aggggagaac agcagcaaac accgcagtag ctgacagcca
60
tttcatagaa gacaaacctа ttgaatatgt acgaactcga taggcgttag ataaatcaga
120
aaagtaaatt aggcaaggta ttccttgcct aatttactag gaagatacag aggtttttaa
180
aaatagttca cctccaaggt tatattagaa gccaacatag cctccgcagg ataagctatt
240
ggagaggttc ttcccttcac ggttgatagt atatgccgtg atgtataagg aacatgcttt
300
atagcgaaga atcaattgat ttgccaatgt gtgtgcagag gttggaactg aacagccatt
360
ggtagtccaa gcaaaatcat tcatagccag gtgagcaaga gtggatgggt tttctcgttt
420
aatatctcga gagtaaccga tggatcccct acctagtaga gcaaaagagc tattcccctaa
480
gcagaattca taaccaatcc cctatgggat ggcgatgttt agaagttttg tcttggaaaa
540

gatcttggaa gtttatattt ccttcaacat atagagcccc accaccttgt ttactgtagt
1037640
tagaattaaa aagactctgt ttagaagatt gaattgagag gtcaccttta gcatagatag
1037700
ctcctcccct ttcttcagca ccgttagttt gtacttttaa tgatgtaata ttctcgaaag
1037760
aaacatcttt ctcagcaaag atcgctcctc catctttagt agctttattt attgagaata
1037820
gtacctttтg agagtctgcg atagtaaggt cgcctttagc atagatagct ccccagatc
1037880
ctggcttatg agagggagtt gatgtttctg tatcggtttg agcagcagga gtagctgtag
1037940
cacaaataaa gtgactttga agattagctg ctgcgggttc agctttactg gaactggagg
1038000
aagacgccga atcatttccg ctagaactac tagaacccga cgtttcgcct gctgtttgag
1038060
cttcaggttt tgtaggtttc ttaacaggag caggaacttc tgcagagttg gaggagaaag
1038120
tcgctttagt aatccctgat agggagattg tagatcctcc aaaaatcgct cctccggata
1038180
gctgggataa gttatttтgg atggttagac ctgtcagatc tgtaaatagc agctctcctt
1038240
gagagaagat agcaccgcct tcaccagtca tttttatctc agacagagtt aaggaacctt
1038300
ctgttcctga tcgtgtcata aaggacaaaa ctccggagtg cgcgttataa aangcgccgc
1038360
caccttтagg atcaggagtc gttgttgtag agtcactatc cttagaaacg gaagcagtct
1038420
ctctgttaga gctagaggag tttgagttcg tgggagttgt tgtattggta gtaggaatgt
1038480
tggtaaattt tgtgaaagat gcgttactag agacgatata ttctgcccca gcttctccaa
1038540
ctgtttcagt aaaagtggta gaggaagaag ttcctacgcg agagaaattt aantnnnnta
1038600
ggttctgg
1038608

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: yes

(viii) POSITION IN THE GENOME: complement (208..501)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Gly Ile Gly Tyr Glu Phe Cys Leu Gly Asn Ser Ser Phe Ala Leu Leu
1               5                  10                  15
Gly Lys Gly Ser Ile Gly Tyr Ser Arg Asp Ile Lys Arg Glu Asn Pro
            20                  25                  30
Ser Thr Leu Ala His Leu Ala Met Asn Asp Phe Ala Trp Thr Thr Asn
        35                  40                  45
Gly Cys Ser Val Pro Thr Ser Ala His Thr Leu Ala Asn Gln Leu Ile
    50                  55                  60
```

CLAIMS

1. Nucleotide sequence having the sequence SEQ ID No. 1 of the *Chlamydia trachomatis LGV2* genome.

2. Nucleotide sequence of *Chlamydia trachomatis*, characterized in that it is chosen from:

a) a nucleotide sequence exhibiting at least 99.9% identity with the sequence SEQ ID No. 1;

b) a nucleotide sequence homologous to the sequence SEQ ID No. 1;

c) a nucleotide sequence complementary to the sequence SEQ ID No. 1 or complementary to a nucleotide sequence as defined in a) or b), and a nucleotide sequence of their corresponding RNA;

d) a nucleotide sequence of a representative fragment of the sequence SEQ ID No. 1, or of a representative fragment of the nucleotide sequence as defined in a), b) or c);

e) a nucleotide sequence comprising a sequence as defined in a), b), c) or d);

f) a nucleotide sequence capable of being obtained from a nucleotide sequence as defined in a), b), c), d) or e); and

g) a modified nucleotide sequence of a nucleotide sequence as defined in a), b), c), d), e) or f).

3. Nucleotide sequence according to Claim 2, characterized in that it is chosen from the ORF2 to ORF1076 sequences.

4. Nucleotide sequences, characterized in that it comprises a nucleotide sequence chosen from:

a) a nucleotide sequence according to Claim 3;

b) a homologous nucleotide sequence exhibiting at least 80% identity with a nucleotide sequence according to Claim 3 or as defined in a);

c) a complementary nucleotide sequence RNA corresponding to a sequence according to Claim 3 or as defined in a) or b);

d) a nucleotide sequence of a representative fragment of a sequence according to Claim 3 or of a sequence as

defined in a), b) or c);

e) a nucleotide sequence capable of being obtained from a sequence according to Claim 3 or as defined in a), b), c) or d); and

f) a modified nucleotide sequence of a sequence according to Claim 3 or as defined in a), b), c), d) or e).

5. Polypeptide encoded by a nucleotide sequence according to one of Claims 2 to 4.

6. Polypeptide according to Claim 5, characterized in that it is encoded by a representative fragment corresponding to an ORF sequence having a nucleotide sequence according to one of Claims 1 to 4.

7. *Chlamydia trachomatis* polypeptide, characterized in that it is chosen from the sequences SEQ ID No. 2 to SEQ ID No. 1076.

8. Polypeptide, characterized in that it comprises a polypeptide chosen from:

a) a polypeptide according to one of Claims 5 to 7;

b) a polypeptide homologous to a polypeptide according to one of Claims 5 to 7, or as defined in a);

c) a fragment of at least 5 amino acids of a polypeptide according to one of Claims 5 to 7, or as defined in a) or b);

d) a biologically active fragment of a polypeptide according to one of Claims 5 to 7, or as defined in a), b) or c); and

e) a modified polypeptide of a polypeptide according to one of Claims 5 to 7, or as defined in a), b), c) or d).

9. Nucleotide sequence encoding a polypeptide according to Claim 8.

10. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a polypeptide of the cellular envelope of *Chlamydia trachomatis* or one of its fragments.

11. Nucleotide sequence according to Claim 10, characterized in that it encodes a polypeptide of the outer cellular envelope of *Chlamydia trachomatis* or one

of its fragments.

12. Nucleotide sequence according to either of Claims 10 and 11, characterized in that it is chosen from the following sequences:

ORF3; ORF19; ORF51; ORF189; ORF212; ORF213; ORF324; ORF477; ORF478; ORF479; ORF481; ORF482; ORF483; ORF484; ORF486; ORF488; ORF489; ORF490; ORF572; ORF573; ORF742; ORF817; ORF818; ORF820; ORF1035; ORF1036; ORF1037; ORF1038; ORF1070; ORF1071; ORF1073 and one of their representative fragments.

13. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 1 and 3 transmembrane domains, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF2; ORF3; ORF5; ORF8; ORF9; ORF10; ORF11; ORF12; ORF17; ORF21; ORF26; ORF27; ORF28; ORF29; ORF30; ORF31; ORF33; ORF35; ORF37; ORF39; ORF40; ORF41; ORF42; ORF43; ORF44; ORF45; ORF46; ORF47; ORF48; ORF49; ORF52; ORF53; ORF55; ORF56; ORF58; ORF65; ORF66; ORF68; ORF70; ORF74; ORF75; ORF76; ORF78; ORF79; ORF81; ORF82; ORF83; ORF86; ORF91; ORF92; ORF94; ORF97; ORF100; ORF102; ORF103; ORF105; ORF106; ORF107; ORF109; ORF110; ORF111; ORF112; ORF113; ORF114; ORF115; ORF116; ORF117; ORF120; ORF122; ORF123; ORF130; ORF134; ORF135; ORF137; ORF140; ORF141; ORF143; ORF144; ORF145; ORF147; ORF148; ORF149; ORF150; ORF151; ORF155; ORF156; ORF162; ORF163; ORF164; ORF165; ORF166; ORF167; ORF168; ORF169; ORF170; ORF171; ORF173; ORF175; ORF176; ORF177; ORF181; ORF183; ORF184; ORF186; ORF187; ORF188; ORF190; ORF191; ORF192; ORF194; ORF195; ORF196; ORF197; ORF198; ORF199; ORF201; ORF202; ORF204; ORF206; ORF207; ORF209; ORF212; ORF213; ORF217; ORF219; ORF220; ORF221; ORF222; ORF223; ORF224; ORF225; ORF227; ORF228; ORF231; ORF232; ORF234; ORF236; ORF237; ORF243; ORF244; ORF245; ORF247; ORF248; ORF249; ORF252; ORF254; ORF257; ORF260; ORF261; ORF263; ORF265; ORF266; ORF267; ORF270; ORF271; ORF272; ORF274; ORF276; ORF277; ORF278;

ORF279; ORF282; ORF283; ORF284; ORF285; ORF287; ORF289;
ORF290; ORF291; ORF294; ORF298; ORF305; ORF306; ORF310;
ORF311; ORF313; ORF315; ORF316; ORF319; ORF320; ORF322;
ORF323; ORF325; ORF326; ORF327; ORF328; ORF330; ORF331;
ORF332; ORF333; ORF334; ORF335; ORF336; ORF338; ORF339;
ORF340; ORF341; ORF344; ORF345; ORF348; ORF349; ORF350;
ORF351; ORF352; ORF353; ORF356; ORF357; ORF358; ORF361;
ORF362; ORF366; ORF367; ORF368; ORF370; ORF372; ORF373;
ORF375; ORF377; ORF378; ORF379; ORF380; ORF382; ORF383;
ORF384; ORF385; ORF387; ORF389; ORF390; ORF391; ORF393;
ORF396; ORF398; ORF399; ORF403; ORF404; ORF406; ORF407;
ORF413; ORF414; ORF417; ORF418; ORF420; ORF421; ORF424;
ORF426; ORF427; ORF428; ORF430; ORF433; ORF434; ORF435;
ORF436; ORF437; ORF440; ORF443; ORF446; ORF448; ORF450;
ORF451; ORF454; ORF455; ORF457; ORF458; ORF459; ORF463;
ORF464; ORF466; ORF467; ORF468; ORF469; ORF470; ORF473;
ORF474; ORF475; ORF476; ORF477; ORF479; ORF480; ORF481;
ORF483; ORF484; ORF485; ORF486; ORF487; ORF488; ORF491;
ORF493; ORF496; ORF497; ORF498; ORF500; ORF501; ORF503;
ORF504; ORF508; ORF512; ORF513; ORF514; ORF519; ORF521;
ORF523; ORF524; ORF526; ORF527; ORF529; ORF530; ORF531;
ORF532; ORF534; ORF536; ORF537; ORF538; ORF540; ORF541;
ORF542; ORF543; ORF544; ORF545; ORF546; ORF547; ORF551;
ORF552; ORF553; ORF555; ORF558; ORF559; ORF560; ORF561;
ORF562; ORF566; ORF567; ORF568; ORF569; ORF571; ORF572;
ORF574; ORF575; ORF576; ORF580; ORF582; ORF585; ORF587;
ORF589; ORF592; ORF593; ORF595; ORF596; ORF597; ORF599;
ORF601; ORF602; ORF603; ORF604; ORF608; ORF609; ORF610;
ORF611; ORF615; ORF616; ORF617; ORF618; ORF621; ORF622;
ORF623; ORF624; ORF625; ORF628; ORF632; ORF633; ORF634;
ORF635; ORF637; ORF638; ORF640; ORF641; ORF643; ORF646;
ORF648; ORF649; ORF651; ORF652; ORF653; ORF654; ORF655;
ORF658; ORF664; ORF665; ORF666; ORF668; ORF669; ORF670;
ORF671; ORF672; ORF673; ORF674; ORF676; ORF677; ORF678;
ORF680; ORF682; ORF683; ORF684; ORF686; ORF688; ORF689;
ORF690; ORF691; ORF692; ORF693; ORF695; ORF696; ORF698;
ORF701; ORF703; ORF704; ORF705; ORF706; ORF707; ORF709;
ORF710; ORF711; ORF712; ORF713; ORF714; ORF715; ORF717;
ORF718; ORF720; ORF721; ORF722; ORF724; ORF726; ORF728;

ORF729; ORF730; ORF731; ORF732; ORF733; ORF734; ORF737; ORF738; ORF739; ORF740; ORF742; ORF743; ORF744; ORF745; ORF746; ORF748; ORF750; ORF751; ORF752; ORF753; ORF754; ORF755; ORF757; ORF758; ORF759; ORF760; ORF764; ORF766; ORF768; ORF769; ORF771; ORF772; ORF773; ORF774; ORF775; ORF776; ORF777; ORF778; ORF779; ORF780; ORF781; ORF782; ORF783; ORF786; ORF787; ORF788; ORF789; ORF790; ORF793; ORF798; ORF800; ORF802; ORF803; ORF806; ORF808; ORF809; ORF810; ORF811; ORF813; ORF814; ORF817; ORF820; ORF822; ORF824; ORF825; ORF827; ORF828; ORF829; ORF830; ORF833; ORF834; ORF835; ORF837; ORF838; ORF839; ORF840; ORF841; ORF842; ORF843; ORF845; ORF848; ORF849; ORF850; ORF851; ORF852; ORF854; ORF855; ORF856; ORF857; ORF859; ORF860; ORF862; ORF863; ORF864; ORF866; ORF869; ORF872; ORF873; ORF874; ORF878; ORF879; ORF880; ORF881; ORF883; ORF884; ORF885; ORF886; ORF887; ORF892; ORF893; ORF894; ORF895; ORF897; ORF899; ORF900; ORF901; ORF904; ORF906; ORF909; ORF910; ORF912; ORF914; ORF917; ORF920; ORF921; ORF922; ORF923; ORF924; ORF925; ORF926; ORF927; ORF930; ORF933; ORF934; ORF935; ORF936; ORF937; ORF940; ORF941; ORF942; ORF943; ORF944; ORF945; ORF947; ORF948; ORF951; ORF952; ORF953; ORF954; ORF955; ORF956; ORF957; ORF958; ORF960; ORF961; ORF962; ORF963; ORF964; ORF966; ORF967; ORF969; ORF970; ORF971; ORF973; ORF974; ORF979; ORF980; ORF981; ORF982; ORF984; ORF988; ORF989; ORF990; ORF991; ORF995; ORF996; ORF999; ORF1001; ORF1003; ORF1004; ORF1005; ORF1006; ORF1007; ORF1009; ORF1010; ORF1011; ORF1012; ORF1013; ORF1014; ORF1016; ORF1017; ORF1018; ORF1020; ORF1021; ORF1025; ORF1026; ORF1027; ORF1029; ORF1030; ORF1031; ORF1035; ORF1036; ORF1037; ORF1038; ORF1039; ORF1040; ORF1044; ORF1045; ORF1047; ORF1048; ORF1050; ORF1051; ORF1052; ORF1053; ORF1055; ORF1056; ORF1057; ORF1058; ORF1061; ORF1062; ORF1063; ORF1064; ORF1065; ORF1066; ORF1068; ORF1069; ORF1072; ORF1074; ORF1076 and one of their representative fragments.

14. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 4 and 6

transmembrane domains and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF7; ORF14; ORF16; ORF32; ORF34; ORF36; ORF38; ORF50; ORF57; ORF59; ORF61; ORF62; ORF63; ORF64; ORF67; ORF69; ORF72; ORF77; ORF80; ORF84; ORF87; ORF93; ORF95; ORF99; ORF108; ORF119; ORF125; ORF126; ORF129; ORF131; ORF136; ORF139; ORF146; ORF152; ORF154; ORF160; ORF161; ORF172; ORF179; ORF182; ORF185; ORF200; ORF203; ORF205; ORF239; ORF242; ORF250; ORF253; ORF256; ORF259; ORF262; ORF268; ORF275; ORF281; ORF286; ORF288; ORF292; ORF295; ORF296; ORF297; ORF299; ORF300; ORF308; ORF314; ORF317; ORF318; ORF324; ORF342; ORF343; ORF355; ORF360; ORF374; ORF376; ORF386; ORF388; ORF392; ORF394; ORF395; ORF402; ORF405; ORF411; ORF415; ORF416; ORF422; ORF423; ORF429; ORF432; ORF441; ORF442; ORF444; ORF449; ORF452; ORF456; ORF460; ORF461; ORF465; ORF471; ORF472; ORF482; ORF489; ORF492; ORF494; ORF495; ORF502; ORF505; ORF506; ORF509; ORF516; ORF517; ORF520; ORF525; ORF533; ORF539; ORF549; ORF554; ORF557; ORF563; ORF570; ORF573; ORF581; ORF590; ORF591; ORF600; ORF607; ORF612; ORF613; ORF620; ORF626; ORF629; ORF630; ORF639; ORF644; ORF647; ORF656; ORF659; ORF661; ORF685; ORF687; ORF699; ORF700; ORF708; ORF716; ORF719; ORF725; ORF747; ORF749; ORF756; ORF765; ORF767; ORF794; ORF796; ORF797; ORF799; ORF801; ORF807; ORF821; ORF823; ORF826; ORF847; ORF853; ORF861; ORF870; ORF871; ORF875; ORF882; ORF888; ORF889; ORF898; ORF902; ORF903; ORF911; ORF916; ORF931; ORF939; ORF975; ORF976; ORF978; ORF983; ORF986; ORF987; ORF992; ORF993; ORF1000; ORF1002; ORF1008; ORF1019; ORF1022; ORF1032; ORF1034; ORF1046; ORF1054; ORF1060; ORF1071 and one of their representative fragments.

15. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having at least 7 transmembrane domains and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF4; ORF6; ORF13; ORF20; ORF51; ORF71; ORF88; ORF118;

ORF128; ORF132; ORF133; ORF158; ORF159; ORF174; ORF180; ORF189; ORF210; ORF211; ORF214; ORF215; ORF226; ORF229; ORF233; ORF235; ORF240; ORF246; ORF251; ORF255; ORF273; ORF354; ORF364; ORF369; ORF371; ORF397; ORF401; ORF409; ORF412; ORF419; ORF439; ORF453; ORF462; ORF490; ORF510; ORF511; ORF518; ORF535; ORF548; ORF550; ORF564; ORF565; ORF578; ORF579; ORF614; ORF631; ORF636; ORF650; ORF662; ORF667; ORF679; ORF681; ORF702; ORF727; ORF741; ORF763; ORF791; ORF792; ORF815; ORF816; ORF832; ORF846; ORF858; ORF865; ORF867; ORF868; ORF877; ORF891; ORF896; ORF907; ORF908; ORF918; ORF919; ORF932; ORF959; ORF977; ORF994; ORF998; ORF1024; ORF1028; ORF1042; ORF1067; ORF1070; ORF1073 and one of their representative fragments.

16. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF10; ORF44; ORF45; ORF46; ORF47; ORF93; ORF101; ORF102; ORF103; ORF106; ORF107; ORF120; ORF121; ORF130; ORF135; ORF140; ORF143; ORF144; ORF145; ORF158; ORF159; ORF160; ORF161; ORF192; ORF193; ORF196; ORF196; ORF197; ORF198; ORF199; ORF227; ORF229; ORF236; ORF236; ORF239; ORF243; ORF245; ORF264; ORF265; ORF297; ORF331; ORF333; ORF359; ORF360; ORF374; ORF404; ORF405; ORF405; ORF410; ORF415; ORF415; ORF416; ORF417; ORF432; ORF460; ORF461; ORF462; ORF495; ORF513; ORF515; ORF566; ORF566; ORF566; ORF589; ORF613; ORF645; ORF646; ORF647; ORF652; ORF653; ORF654; ORF672; ORF673; ORF674; ORF682; ORF684; ORF692; ORF700; ORF725; ORF801; ORF802; ORF835; ORF836; ORF837; ORF860; ORF861; ORF862; ORF863; ORF869; ORF869; ORF925; ORF964; ORF983 and one of their representative fragments.

17. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its

fragments which is involved in the metabolism of nucleotides, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF142; ORF142; ORF169; ORF256; ORF268; ORF325; ORF352; ORF366; ORF435; ORF444; ORF528; ORF529; ORF530; ORF548; ORF549; ORF601; ORF602; ORF617; ORF619; ORF644; ORF745; ORF971; ORF972; ORF1023 and one of their representative fragments.

18. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleic acids, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF5; ORF12; ORF82; ORF96; ORF97; ORF98; ORF99; ORF100; ORF105; ORF118; ORF136; ORF137; ORF163; ORF190; ORF204; ORF259; ORF260; ORF262; ORF290; ORF300; ORF301; ORF302; ORF387; ORF427; ORF434; ORF441; ORF444; ORF471; ORF595; ORF596; ORF597; ORF599; ORF600; ORF605; ORF612; ORF624; ORF625; ORF650; ORF657; ORF658; ORF702; ORF703; ORF704; ORF708; ORF719; ORF766; ORF767; ORF775; ORF779; ORF787; ORF788; ORF794; ORF841; ORF842; ORF883; ORF884; ORF907; ORF918; ORF924; ORF928; ORF929; ORF962; ORF962; ORF963; ORF969; ORF970; ORF975; ORF979; ORF995; ORF1031; ORF1032 and one of their representative fragments.

19. Nucleotide sequence according to one of Claims 2 to 5, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of amino acids, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF27; ORF41; ORF55; ORF56; ORF57; ORF59; ORF62; ORF63; ORF64; ORF65; ORF119; ORF132; ORF240; ORF241; ORF277; ORF278; ORF279; ORF382; ORF406; ORF428; ORF442; ORF446; ORF447; ORF453; ORF454; ORF541; ORF542; ORF591; ORF608; ORF609; ORF610; ORF618; ORF648; ORF649; ORF660; ORF661; ORF677; ORF717; ORF765; ORF797; ORF871; ORF875; ORF920; ORF922; ORF937; ORF998; ORF1020; ORF1021; ORF1034; ORF1044; ORF1046; ORF1049 and one of their

representative fragments.

20. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of polypeptides, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF21; ORF21; ORF22; ORF23; ORF24; ORF25; ORF26; ORF75; ORF84; ORF84; ORF86; ORF92; ORF133; ORF151; ORF152; ORF157; ORF179; ORF209; ORF307; ORF326; ORF343; ORF344; ORF345; ORF371; ORF429; ORF519; ORF557; ORF586; ORF587; ORF630; ORF656; ORF706; ORF707; ORF730; ORF751; ORF752; ORF786; ORF847; ORF885; ORF923; ORF978; ORF1039; ORF1048 and one of their representative fragments.

21. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of fatty acids, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF4; ORF15; ORF16; ORF141; ORF173; ORF205; ORF205; ORF206; ORF207; ORF208; ORF312; ORF355; ORF415; ORF550; ORF558; ORF560; ORF561; ORF574; ORF574; ORF577; ORF578; ORF590; ORF614; ORF772; ORF808; ORF809; ORF904; ORF905; ORF905; ORF933; ORF934; ORF934; ORF936 and one of their representative fragments.

22. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the synthesis of the wall, and in that it comprises a nucleotide sequence chosen from the following sequences:
ORF87; ORF196; ORF242; ORF269; ORF628; ORF629; ORF634; ORF635; ORF637; ORF638; ORF1019 and one of their representative fragments.

23. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the transcription,

translation and/or maturation process, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF112; ORF113; ORF332; ORF212; ORF213; ORF350; ORF362; ORF363; ORF364; ORF407; ORF451; ORF546; ORF643; ORF744; ORF746; ORF833; ORF868; ORF981; ORF982; ORF1003; ORF1011; ORF1042 and one of their representative fragments.

24. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* ribosomal polypeptide or one of its fragments, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF114; ORF115; ORF116; ORF328; ORF361; ORF375; ORF445; ORF543; ORF584; ORF585; ORF743; ORF813; ORF941; ORF942; ORF944; ORF946; ORF947; ORF948; ORF950; ORF951; ORF952; ORF953; ORF954; ORF955; ORF955; ORF957; ORF958; ORF960; ORF961; ORF1040; ORF1041; ORF1043; ORF1063; ORF1064 and one of their representative fragments.

25. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* transport polypeptide or one of its fragments, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF6; ORF50; ORF51; ORF80; ORF125; ORF126; ORF128; ORF129; ORF215; ORF246; ORF248; ORF249; ORF251; ORF252; ORF253; ORF255; ORF271; ORF275; ORF293; ORF309; ORF323; ORF324; ORF398; ORF401; ORF449; ORF511; ORF512; ORF564; ORF565; ORF667; ORF679; ORF680; ORF711; ORF712; ORF713; ORF714; ORF715; ORF730; ORF731; ORF736; ORF737; ORF738; ORF870; ORF908; ORF919; ORF977; ORF987; ORF988; ORF992; ORF993; ORF994; ORF1028; ORF1029 and one of their representative fragments.

26. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the virulence process, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF20; ORF815; ORF816; ORF898; ORF1059; ORF1060 and one of their representative fragments.

27. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the secretory system and/or which is secreted, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF758; ORF888; ORF889; ORF890; ORF891; ORF896; ORF897; ORF898 and one of their representative fragments.

28. Nucleotide sequence according to one of Claims 2 to 4, and 9, characterized in that it encodes a polypeptide specific to *Chlamydiae* or one of its fragments, and in that it comprises a nucleotide sequence chosen from the following sequences:

ORF22; ORF29; ORF31; ORF32; ORF34; ORF35; ORF39; ORF40; ORF43; ORF48; ORF49; ORF50; ORF52; ORF53; ORF54; ORF72; ORF77; ORF78; ORF87; ORF90; ORF95; ORF108; ORF110; ORF111; ORF122; ORF123; ORF124; ORF127; ORF138; ORF144; ORF146; ORF153; ORF155; ORF164; ORF166; ORF175; ORF182; ORF184; ORF186; ORF187; ORF188; ORF202; ORF210; ORF247; ORF258; ORF266; ORF267; ORF270; ORF273; ORF274; ORF295; ORF296; ORF305; ORF306; ORF309; ORF318; ORF319; ORF322; ORF326; ORF342; ORF357; ORF376; ORF379; ORF380; ORF388; ORF390; ORF400; ORF431; ORF433; ORF438; ORF443; ORF456; ORF457; ORF458; ORF464; ORF468; ORF470; ORF473; ORF486; ORF489; ORF497; ORF501; ORF503; ORF504; ORF508; ORF512; ORF521; ORF522; ORF523; ORF524; ORF533; ORF535; ORF536; ORF537; ORF538; ORF539; ORF540; ORF554; ORF563; ORF572; ORF579; ORF595; ORF603; ORF604; ORF606; ORF607; ORF615; ORF616; ORF622; ORF641; ORF642; ORF659; ORF668; ORF670; ORF693; ORF695; ORF696; ORF699; ORF703; ORF704; ORF716; ORF726; ORF728; ORF739; ORF742; ORF747; ORF750; ORF751; ORF755; ORF757; ORF759; ORF761; ORF762; ORF763; ORF764; ORF773; ORF780; ORF781; ORF789; ORF800; ORF803; ORF804; ORF818; ORF820; ORF822; ORF823; ORF824; ORF827; ORF828; ORF839; ORF849; ORF850; ORF851; ORF852; ORF855; ORF856; ORF857; ORF858; ORF859; ORF860; ORF861; ORF862; ORF863;

ORF865; ORF868; ORF869; ORF870; ORF871; ORF872; ORF873; ORF874; ORF875; ORF877; ORF878; ORF880; ORF882; ORF884; ORF886; ORF893; ORF901; ORF906; ORF910; ORF912; ORF915; ORF916; ORF917; ORF926; ORF929; ORF933; ORF965; ORF967; ORF968; ORF984; ORF986; ORF989; ORF990; ORF996; ORF997; ORF1001; ORF1002; ORF1013; ORF1016; ORF1031; ORF1033; ORF1035; ORF1049; ORF1051; ORF1052; ORF1054; ORF1056; ORF1057; ORF1058; ORF1062; ORF1070; ORF1071; ORF1073 and one of their representative fragments.

29. Polypeptide according to one of Claims 5 to 8, characterized in that it is a polypeptide of the cellular envelope of *Chlamydia trachomatis* or one of its fragments.

30. Polypeptide according to Claim 29, characterized in that it is a polypeptide of the outer cellular envelope of *Chlamydia trachomatis* or one of its fragments.

31. Polypeptide according to either of Claims 29 and 30, characterized in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 3; SEQ ID No. 19; SEQ ID No. 51; SEQ ID No. 189; SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 324; SEQ ID No. 477; SEQ ID No. 478; SEQ ID No. 479; SEQ ID No. 481; SEQ ID No. 482; SEQ ID No. 483; SEQ ID No. 484; SEQ ID No. 486; SEQ ID No. 488; SEQ ID No. 489; SEQ ID No. 490; SEQ ID No. 572; SEQ ID No. 573; SEQ ID No. 742; SEQ ID No. 817; SEQ ID No. 818; SEQ ID No. 820; SEQ ID No. 1035; SEQ ID No. 1036; SEQ ID No. 1037; SEQ ID No. 1038; SEQ ID No. 1070; SEQ ID No. 1071; SEQ ID No. 1073 and one of their fragments.

32. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 1 and 3 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 2; SEQ ID No. 3; SEQ ID No. 5; SEQ ID No. 8; SEQ ID No. 9; SEQ ID No. 10; SEQ ID No. 11;

SEQ ID No. 12; SEQ ID No. 17; SEQ ID No. 21;
SEQ ID No. 26; SEQ ID No. 27; SEQ ID No. 28;
SEQ ID No. 29; SEQ ID No. 30; SEQ ID No. 31;
SEQ ID No. 33; SEQ ID No. 35; SEQ ID No. 37;
SEQ ID No. 39; SEQ ID No. 40; SEQ ID No. 41;
SEQ ID No. 42; SEQ ID No. 43; SEQ ID No. 44;
SEQ ID No. 45; SEQ ID No. 46; SEQ ID No. 47;
SEQ ID No. 48; SEQ ID No. 49; SEQ ID No. 52;
SEQ ID No. 53; SEQ ID No. 55; SEQ ID No. 56;
SEQ ID No. 58; SEQ ID No. 65; SEQ ID No. 66;
SEQ ID No. 68; SEQ ID No. 70; SEQ ID No. 74;
SEQ ID No. 75; SEQ ID No. 76; SEQ ID No. 78;
SEQ ID No. 79; SEQ ID No. 81; SEQ ID No. 82;
SEQ ID No. 83; SEQ ID No. 86; SEQ ID No. 91;
SEQ ID No. 92; SEQ ID No. 94; SEQ ID No. 97;
SEQ ID No. 100; SEQ ID No. 102; SEQ ID No. 103;
SEQ ID No. 105; SEQ ID No. 106; SEQ ID No. 107;
SEQ ID No. 109; SEQ ID No. 110; SEQ ID No. 111;
SEQ ID No. 112; SEQ ID No. 113; SEQ ID No. 114;
SEQ ID No. 115; SEQ ID No. 116; SEQ ID No. 117;
SEQ ID No. 120; SEQ ID No. 122; SEQ ID No. 123;
SEQ ID No. 130; SEQ ID No. 134; SEQ ID No. 135;
SEQ ID No. 137; SEQ ID No. 140; SEQ ID No. 141;
SEQ ID No. 143; SEQ ID No. 144; SEQ ID No. 145;
SEQ ID No. 147; SEQ ID No. 148; SEQ ID No. 149;
SEQ ID No. 150; SEQ ID No. 151; SEQ ID No. 155;
SEQ ID No. 156; SEQ ID No. 162; SEQ ID No. 163;
SEQ ID No. 164; SEQ ID No. 165; SEQ ID No. 166;
SEQ ID No. 167; SEQ ID No. 168; SEQ ID No. 169;
SEQ ID No. 170; SEQ ID No. 171; SEQ ID No. 173;
SEQ ID No. 175; SEQ ID No. 176; SEQ ID No. 177;
SEQ ID No. 181; SEQ ID No. 183; SEQ ID No. 184;
SEQ ID No. 186; SEQ ID No. 187; SEQ ID No. 188;
SEQ ID No. 190; SEQ ID No. 191; SEQ ID No. 192;
SEQ ID No. 194; SEQ ID No. 195; SEQ ID No. 196;
SEQ ID No. 197; SEQ ID No. 198; SEQ ID No. 199;
SEQ ID No. 201; SEQ ID No. 202; SEQ ID No. 204;
SEQ ID No. 206; SEQ ID No. 207; SEQ ID No. 209;
SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 217;

SEQ ID No. 219; SEQ ID No. 220; SEQ ID No. 221;
SEQ ID No. 222; SEQ ID No. 223; SEQ ID No. 224;
SEQ ID No. 225; SEQ ID No. 227; SEQ ID No. 228;
SEQ ID No. 231; SEQ ID No. 232; SEQ ID No. 234;
SEQ ID No. 236; SEQ ID No. 237; SEQ ID No. 243;
SEQ ID No. 244; SEQ ID No. 245; SEQ ID No. 247;
SEQ ID No. 248; SEQ ID No. 249; SEQ ID No. 252;
SEQ ID No. 254; SEQ ID No. 257; SEQ ID No. 260;
SEQ ID No. 261; SEQ ID No. 263; SEQ ID No. 265;
SEQ ID No. 266; SEQ ID No. 267; SEQ ID No. 270;
SEQ ID No. 271; SEQ ID No. 272; SEQ ID No. 274;
SEQ ID No. 276; SEQ ID No. 277; SEQ ID No. 278;
SEQ ID No. 279; SEQ ID No. 282; SEQ ID No. 283;
SEQ ID No. 284; SEQ ID No. 285; SEQ ID No. 287;
SEQ ID No. 289; SEQ ID No. 290; SEQ ID No. 291;
SEQ ID No. 294; SEQ ID No. 298; SEQ ID No. 305;
SEQ ID No. 306; SEQ ID No. 310; SEQ ID No. 311;
SEQ ID No. 313; SEQ ID No. 315; SEQ ID No. 316;
SEQ ID No. 319; SEQ ID No. 320; SEQ ID No. 322;
SEQ ID No. 323; SEQ ID No. 325; SEQ ID No. 326;
SEQ ID No. 327; SEQ ID No. 328; SEQ ID No. 330;
SEQ ID No. 331; SEQ ID No. 332; SEQ ID No. 333;
SEQ ID No. 334; SEQ ID No. 335; SEQ ID No. 336;
SEQ ID No. 338; SEQ ID No. 339; SEQ ID No. 340;
SEQ ID No. 341; SEQ ID No. 344; SEQ ID No. 345;
SEQ ID No. 348; SEQ ID No. 349; SEQ ID No. 350;
SEQ ID No. 351; SEQ ID No. 352; SEQ ID No. 353;
SEQ ID No. 356; SEQ ID No. 357; SEQ ID No. 358;
SEQ ID No. 361; SEQ ID No. 362; SEQ ID No. 366;
SEQ ID No. 367; SEQ ID No. 368; SEQ ID No. 370;
SEQ ID No. 372; SEQ ID No. 373; SEQ ID No. 375;
SEQ ID No. 377; SEQ ID No. 378; SEQ ID No. 379;
SEQ ID No. 380; SEQ ID No. 382; SEQ ID No. 383;
SEQ ID No. 384; SEQ ID No. 385; SEQ ID No. 387;
SEQ ID No. 389; SEQ ID No. 390; SEQ ID No. 391;
SEQ ID No. 393; SEQ ID No. 396; SEQ ID No. 398;
SEQ ID No. 399; SEQ ID No. 403; SEQ ID No. 404;
SEQ ID No. 406; SEQ ID No. 407; SEQ ID No. 413;
SEQ ID No. 414; SEQ ID No. 417; SEQ ID No. 418;

SEQ ID No. 420; SEQ ID No. 421; SEQ ID No. 424;
SEQ ID No. 426; SEQ ID No. 427; SEQ ID No. 428;
SEQ ID No. 430; SEQ ID No. 433; SEQ ID No. 434;
SEQ ID No. 435; SEQ ID No. 436; SEQ ID No. 437;
SEQ ID No. 440; SEQ ID No. 443; SEQ ID No. 446;
SEQ ID No. 448; SEQ ID No. 450; SEQ ID No. 451;
SEQ ID No. 454; SEQ ID No. 455; SEQ ID No. 457;
SEQ ID No. 458; SEQ ID No. 459; SEQ ID No. 463;
SEQ ID No. 464; SEQ ID No. 466; SEQ ID No. 467;
SEQ ID No. 468; SEQ ID No. 469; SEQ ID No. 470;
SEQ ID No. 473; SEQ ID No. 474; SEQ ID No. 475;
SEQ ID No. 476; SEQ ID No. 477; SEQ ID No. 479;
SEQ ID No. 480; SEQ ID No. 481; SEQ ID No. 483;
SEQ ID No. 484; SEQ ID No. 485; SEQ ID No. 486;
SEQ ID No. 487; SEQ ID No. 488; SEQ ID No. 491;
SEQ ID No. 493; SEQ ID No. 496; SEQ ID No. 497;
SEQ ID No. 498; SEQ ID No. 500; SEQ ID No. 501;
SEQ ID No. 503; SEQ ID No. 504; SEQ ID No. 508;
SEQ ID No. 512; SEQ ID No. 513; SEQ ID No. 514;
SEQ ID No. 519; SEQ ID No. 521; SEQ ID No. 523;
SEQ ID No. 524; SEQ ID No. 526; SEQ ID No. 527;
SEQ ID No. 529; SEQ ID No. 530; SEQ ID No. 531;
SEQ ID No. 532; SEQ ID No. 534; SEQ ID No. 536;
SEQ ID No. 537; SEQ ID No. 538; SEQ ID No. 540;
SEQ ID No. 541; SEQ ID No. 542; SEQ ID No. 543;
SEQ ID No. 544; SEQ ID No. 545; SEQ ID No. 546;
SEQ ID No. 547; SEQ ID No. 551; SEQ ID No. 552;
SEQ ID No. 553; SEQ ID No. 555; SEQ ID No. 558;
SEQ ID No. 559; SEQ ID No. 560; SEQ ID No. 561;
SEQ ID No. 562; SEQ ID No. 566; SEQ ID No. 567;
SEQ ID No. 568; SEQ ID No. 569; SEQ ID No. 571;
SEQ ID No. 572; SEQ ID No. 574; SEQ ID No. 575;
SEQ ID No. 576; SEQ ID No. 580; SEQ ID No. 582;
SEQ ID No. 585; SEQ ID No. 587; SEQ ID No. 589;
SEQ ID No. 592; SEQ ID No. 593; SEQ ID No. 595;
SEQ ID No. 596; SEQ ID No. 597; SEQ ID No. 599;
SEQ ID No. 601; SEQ ID No. 602; SEQ ID No. 603;
SEQ ID No. 604; SEQ ID No. 608; SEQ ID No. 609;
SEQ ID No. 610; SEQ ID No. 611; SEQ ID No. 615;

SEQ ID No. 616; SEQ ID No. 617; SEQ ID No. 618;
SEQ ID No. 621; SEQ ID No. 622; SEQ ID No. 623;
SEQ ID No. 624; SEQ ID No. 625; SEQ ID No. 628;
SEQ ID No. 632; SEQ ID No. 633; SEQ ID No. 634;
SEQ ID No. 635; SEQ ID No. 637; SEQ ID No. 638;
SEQ ID No. 640; SEQ ID No. 641; SEQ ID No. 643;
SEQ ID No. 646; SEQ ID No. 648; SEQ ID No. 649;
SEQ ID No. 651; SEQ ID No. 652; SEQ ID No. 653;
SEQ ID No. 654; SEQ ID No. 655; SEQ ID No. 658;
SEQ ID No. 664; SEQ ID No. 665; SEQ ID No. 666;
SEQ ID No. 668; SEQ ID No. 669; SEQ ID No. 670;
SEQ ID No. 671; SEQ ID No. 672; SEQ ID No. 673;
SEQ ID No. 674; SEQ ID No. 676; SEQ ID No. 677;
SEQ ID No. 678; SEQ ID No. 680; SEQ ID No. 682;
SEQ ID No. 683; SEQ ID No. 684; SEQ ID No. 686;
SEQ ID No. 688; SEQ ID No. 689; SEQ ID No. 690;
SEQ ID No. 691; SEQ ID No. 692; SEQ ID No. 693;
SEQ ID No. 695; SEQ ID No. 696; SEQ ID No. 698;
SEQ ID No. 701; SEQ ID No. 703; SEQ ID No. 704;
SEQ ID No. 705; SEQ ID No. 706; SEQ ID No. 707;
SEQ ID No. 709; SEQ ID No. 710; SEQ ID No. 711;
SEQ ID No. 712; SEQ ID No. 713; SEQ ID No. 714;
SEQ ID No. 715; SEQ ID No. 717; SEQ ID No. 718;
SEQ ID No. 720; SEQ ID No. 721; SEQ ID No. 722;
SEQ ID No. 724; SEQ ID No. 726; SEQ ID No. 728;
SEQ ID No. 729; SEQ ID No. 730; SEQ ID No. 731;
SEQ ID No. 732; SEQ ID No. 733; SEQ ID No. 734;
SEQ ID No. 737; SEQ ID No. 738; SEQ ID No. 739;
SEQ ID No. 740; SEQ ID No. 742; SEQ ID No. 743;
SEQ ID No. 744; SEQ ID No. 745; SEQ ID No. 746;
SEQ ID No. 748; SEQ ID No. 750; SEQ ID No. 751;
SEQ ID No. 752; SEQ ID No. 753; SEQ ID No. 754;
SEQ ID No. 755; SEQ ID No. 757; SEQ ID No. 758;
SEQ ID No. 759; SEQ ID No. 760; SEQ ID No. 764;
SEQ ID No. 766; SEQ ID No. 768; SEQ ID No. 769;
SEQ ID No. 771; SEQ ID No. 772; SEQ ID No. 773;
SEQ ID No. 774; SEQ ID No. 775; SEQ ID No. 776;
SEQ ID No. 777; SEQ ID No. 778; SEQ ID No. 779;
SEQ ID No. 780; SEQ ID No. 781; SEQ ID No. 782;

SEQ ID No. 783; SEQ ID No. 786; SEQ ID No. 787;
SEQ ID No. 788; SEQ ID No. 789; SEQ ID No. 790;
SEQ ID No. 793; SEQ ID No. 798; SEQ ID No. 800;
SEQ ID No. 802; SEQ ID No. 803; SEQ ID No. 806;
SEQ ID No. 808; SEQ ID No. 809; SEQ ID No. 810;
SEQ ID No. 811; SEQ ID No. 813; SEQ ID No. 814;
SEQ ID No. 817; SEQ ID No. 820; SEQ ID No. 822;
SEQ ID No. 824; SEQ ID No. 825; SEQ ID No. 827;
SEQ ID No. 828; SEQ ID No. 829; SEQ ID No. 830;
SEQ ID No. 833; SEQ ID No. 834; SEQ ID No. 835;
SEQ ID No. 837; SEQ ID No. 838; SEQ ID No. 839;
SEQ ID No. 840; SEQ ID No. 841; SEQ ID No. 842;
SEQ ID No. 843; SEQ ID No. 845; SEQ ID No. 848;
SEQ ID No. 849; SEQ ID No. 850; SEQ ID No. 851;
SEQ ID No. 852; SEQ ID No. 854; SEQ ID No. 855;
SEQ ID No. 856; SEQ ID No. 857; SEQ ID No. 859;
SEQ ID No. 860; SEQ ID No. 862; SEQ ID No. 863;
SEQ ID No. 864; SEQ ID No. 866; SEQ ID No. 869;
SEQ ID No. 872; SEQ ID No. 873; SEQ ID No. 874;
SEQ ID No. 878; SEQ ID No. 879; SEQ ID No. 880;
SEQ ID No. 881; SEQ ID No. 883; SEQ ID No. 884;
SEQ ID No. 885; SEQ ID No. 886; SEQ ID No. 887;
SEQ ID No. 892; SEQ ID No. 893; SEQ ID No. 894;
SEQ ID No. 895; SEQ ID No. 897; SEQ ID No. 899;
SEQ ID No. 900; SEQ ID No. 901; SEQ ID No. 904;
SEQ ID No. 906; SEQ ID No. 909; SEQ ID No. 910;
SEQ ID No. 912; SEQ ID No. 914; SEQ ID No. 917;
SEQ ID No. 920; SEQ ID No. 921; SEQ ID No. 922;
SEQ ID No. 923; SEQ ID No. 924; SEQ ID No. 925;
SEQ ID No. 926; SEQ ID No. 927; SEQ ID No. 930;
SEQ ID No. 933; SEQ ID No. 934; SEQ ID No. 935;
SEQ ID No. 936; SEQ ID No. 937; SEQ ID No. 940;
SEQ ID No. 941; SEQ ID No. 942; SEQ ID No. 943;
SEQ ID No. 944; SEQ ID No. 945; SEQ ID No. 947;
SEQ ID No. 948; SEQ ID No. 951; SEQ ID No. 952;
SEQ ID No. 953; SEQ ID No. 954; SEQ ID No. 955;
SEQ ID No. 956; SEQ ID No. 957; SEQ ID No. 958;
SEQ ID No. 960; SEQ ID No. 961; SEQ ID No. 962;
SEQ ID No. 963; SEQ ID No. 964; SEQ ID No. 966;

SEQ ID No. 967; SEQ ID No. 969; SEQ ID No. 970;
SEQ ID No. 971; SEQ ID No. 973; SEQ ID No. 974;
SEQ ID No. 979; SEQ ID No. 980; SEQ ID No. 981;
SEQ ID No. 982; SEQ ID No. 984; SEQ ID No. 988;
SEQ ID No. 989; SEQ ID No. 990; SEQ ID No. 991;
SEQ ID No. 995; SEQ ID No. 996; SEQ ID No. 999;
SEQ ID No. 1001; SEQ ID No. 1003; SEQ ID No. 1004;
SEQ ID No. 1005; SEQ ID No. 1006; SEQ ID No. 1007;
SEQ ID No. 1009; SEQ ID No. 1010; SEQ ID No. 1011;
SEQ ID No. 1012; SEQ ID No. 1013; SEQ ID No. 1014;
SEQ ID No. 1016; SEQ ID No. 1017; SEQ ID No. 1018;
SEQ ID No. 1020; SEQ ID No. 1021; SEQ ID No. 1025;
SEQ ID No. 1026; SEQ ID No. 1027; SEQ ID No. 1029;
SEQ ID No. 1030; SEQ ID No. 1031; SEQ ID No. 1035;
SEQ ID No. 1036; SEQ ID No. 1037; SEQ ID No. 1038;
SEQ ID No. 1039; SEQ ID No. 1040; SEQ ID No. 1044;
SEQ ID No. 1045; SEQ ID No. 1047; SEQ ID No. 1048;
SEQ ID No. 1050; SEQ ID No. 1051; SEQ ID No. 1052;
SEQ ID No. 1053; SEQ ID No. 1055; SEQ ID No. 1056;
SEQ ID No. 1057; SEQ ID No. 1058; SEQ ID No. 1061;
SEQ ID No. 1062; SEQ ID No. 1063; SEQ ID No. 1064;
SEQ ID No. 1065; SEQ ID No. 1066; SEQ ID No. 1068;
SEQ ID No. 1069; SEQ ID No. 1072; SEQ ID No. 1074;
SEQ ID No. 1076 and one of their fragments.

33. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having between 4 and 6 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 7; SEQ ID No. 14; SEQ ID No. 16;
SEQ ID No. 32; SEQ ID No. 34; SEQ ID No. 36;
SEQ ID No. 38; SEQ ID No. 50; SEQ ID No. 57;
SEQ ID No. 59; SEQ ID No. 61; SEQ ID No. 62;
SEQ ID No. 63; SEQ ID No. 64; SEQ ID No. 67;
SEQ ID No. 69; SEQ ID No. 72; SEQ ID No. 77;
SEQ ID No. 80; SEQ ID No. 84; SEQ ID No. 87;
SEQ ID No. 93; SEQ ID No. 95; SEQ ID No. 99;
SEQ ID No. 108; SEQ ID No. 119; SEQ ID No. 125;

SEQ ID No. 126; SEQ ID No. 129; SEQ ID No. 131;
SEQ ID No. 136; SEQ ID No. 139; SEQ ID No. 146;
SEQ ID No. 152; SEQ ID No. 154; SEQ ID No. 160;
SEQ ID No. 161; SEQ ID No. 172; SEQ ID No. 179;
SEQ ID No. 182; SEQ ID No. 185; SEQ ID No. 200;
SEQ ID No. 203; SEQ ID No. 205; SEQ ID No. 239;
SEQ ID No. 242; SEQ ID No. 250; SEQ ID No. 253;
SEQ ID No. 256; SEQ ID No. 259; SEQ ID No. 262;
SEQ ID No. 268; SEQ ID No. 275; SEQ ID No. 281;
SEQ ID No. 286; SEQ ID No. 288; SEQ ID No. 292;
SEQ ID No. 295; SEQ ID No. 296; SEQ ID No. 297;
SEQ ID No. 299; SEQ ID No. 300; SEQ ID No. 308;
SEQ ID No. 314; SEQ ID No. 317; SEQ ID No. 318;
SEQ ID No. 324; SEQ ID No. 342; SEQ ID No. 343;
SEQ ID No. 355; SEQ ID No. 360; SEQ ID No. 374;
SEQ ID No. 376; SEQ ID No. 386; SEQ ID No. 388;
SEQ ID No. 392; SEQ ID No. 394; SEQ ID No. 395;
SEQ ID No. 402; SEQ ID No. 405; SEQ ID No. 411;
SEQ ID No. 415; SEQ ID No. 416; SEQ ID No. 422;
SEQ ID No. 423; SEQ ID No. 429; SEQ ID No. 432;
SEQ ID No. 441; SEQ ID No. 442; SEQ ID No. 444;
SEQ ID No. 449; SEQ ID No. 452; SEQ ID No. 456;
SEQ ID No. 460; SEQ ID No. 461; SEQ ID No. 465;
SEQ ID No. 471; SEQ ID No. 472; SEQ ID No. 482;
SEQ ID No. 489; SEQ ID No. 492; SEQ ID No. 494;
SEQ ID No. 495; SEQ ID No. 502; SEQ ID No. 505;
SEQ ID No. 506; SEQ ID No. 509; SEQ ID No. 516;
SEQ ID No. 517; SEQ ID No. 520; SEQ ID No. 525;
SEQ ID No. 533; SEQ ID No. 539; SEQ ID No. 549;
SEQ ID No. 554; SEQ ID No. 557; SEQ ID No. 563;
SEQ ID No. 570; SEQ ID No. 573; SEQ ID No. 581;
SEQ ID No. 590; SEQ ID No. 591; SEQ ID No. 600;
SEQ ID No. 607; SEQ ID No. 612; SEQ ID No. 613;
SEQ ID No. 620; SEQ ID No. 626; SEQ ID No. 629;
SEQ ID No. 630; SEQ ID No. 639; SEQ ID No. 644;
SEQ ID No. 647; SEQ ID No. 656; SEQ ID No. 659;
SEQ ID No. 661; SEQ ID No. 685; SEQ ID No. 687;
SEQ ID No. 699; SEQ ID No. 700; SEQ ID No. 708;
SEQ ID No. 716; SEQ ID No. 719; SEQ ID No. 725;

SEQ ID No. 747; SEQ ID No. 749; SEQ ID No. 756;
SEQ ID No. 765; SEQ ID No. 767; SEQ ID No. 794;
SEQ ID No. 796; SEQ ID No. 797; SEQ ID No. 799;
SEQ ID No. 801; SEQ ID No. 807; SEQ ID No. 821;
SEQ ID No. 823; SEQ ID No. 826; SEQ ID No. 847;
SEQ ID No. 853; SEQ ID No. 861; SEQ ID No. 870;
SEQ ID No. 871; SEQ ID No. 875; SEQ ID No. 882;
SEQ ID No. 888; SEQ ID No. 889; SEQ ID No. 898;
SEQ ID No. 902; SEQ ID No. 903; SEQ ID No. 911;
SEQ ID No. 916; SEQ ID No. 931; SEQ ID No. 939;
SEQ ID No. 975; SEQ ID No. 976; SEQ ID No. 978;
SEQ ID No. 983; SEQ ID No. 986; SEQ ID No. 987;
SEQ ID No. 992; SEQ ID No. 993; SEQ ID No. 1000;
SEQ ID No. 1002; SEQ ID No. 1008; SEQ ID No. 1019;
SEQ ID No. 1022; SEQ ID No. 1032; SEQ ID No. 1034;
SEQ ID No. 1046; SEQ ID No. 1054; SEQ ID No. 1060;
SEQ ID No. 1071 and one of their representative fragments.

34. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* transmembrane polypeptide or one of its fragments, having at least 7 transmembrane domains, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 4; SEQ ID No. 6; SEQ ID No. 13;
SEQ ID No. 20; SEQ ID No. 51; SEQ ID No. 71;
SEQ ID No. 88; SEQ ID No. 118; SEQ ID No. 128;
SEQ ID No. 132; SEQ ID No. 133; SEQ ID No. 158;
SEQ ID No. 159; SEQ ID No. 174; SEQ ID No. 180;
SEQ ID No. 189; SEQ ID No. 210; SEQ ID No. 211;
SEQ ID No. 214; SEQ ID No. 215; SEQ ID No. 226;
SEQ ID No. 229; SEQ ID No. 233; SEQ ID No. 235;
SEQ ID No. 240; SEQ ID No. 246; SEQ ID No. 251;
SEQ ID No. 255; SEQ ID No. 273; SEQ ID No. 354;
SEQ ID No. 364; SEQ ID No. 369; SEQ ID No. 371;
SEQ ID No. 397; SEQ ID No. 401; SEQ ID No. 409;
SEQ ID No. 412; SEQ ID No. 419; SEQ ID No. 439;
SEQ ID No. 453; SEQ ID No. 462; SEQ ID No. 490;
SEQ ID No. 510; SEQ ID No. 511; SEQ ID No. 518;

SEQ ID No. 535; SEQ ID No. 548; SEQ ID No. 550;
SEQ ID No. 564; SEQ ID No. 565; SEQ ID No. 578;
SEQ ID No. 579; SEQ ID No. 614; SEQ ID No. 631;
SEQ ID No. 636; SEQ ID No. 650; SEQ ID No. 662;
SEQ ID No. 667; SEQ ID No. 679; SEQ ID No. 681;
SEQ ID No. 702; SEQ ID No. 727; SEQ ID No. 741;
SEQ ID No. 763; SEQ ID No. 791; SEQ ID No. 792;
SEQ ID No. 815; SEQ ID No. 816; SEQ ID No. 832;
SEQ ID No. 846; SEQ ID No. 858; SEQ ID No. 865;
SEQ ID No. 867; SEQ ID No. 868; SEQ ID No. 877;
SEQ ID No. 891; SEQ ID No. 896; SEQ ID No. 907;
SEQ ID No. 908; SEQ ID No. 918; SEQ ID No. 919;
SEQ ID No. 932; SEQ ID No. 959; SEQ ID No. 977;
SEQ ID No. 994; SEQ ID No. 998; SEQ ID No. 1024;
SEQ ID No. 1028; SEQ ID No. 1042; SEQ ID No. 1067;
SEQ ID No. 1070; SEQ ID No. 1073 and one of their representative fragments.

35. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the intermediate metabolism, in particular in the metabolism of sugars and/or of cofactors, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 10; SEQ ID No. 44; SEQ ID No. 45;
SEQ ID No. 46; SEQ ID No. 47; SEQ ID No. 93;
SEQ ID No. 101; SEQ ID No. 102; SEQ ID No. 103;
SEQ ID No. 106; SEQ ID No. 107; SEQ ID No. 120;
SEQ ID No. 121; SEQ ID No. 130; SEQ ID No. 135;
SEQ ID No. 140; SEQ ID No. 143; SEQ ID No. 144;
SEQ ID No. 145; SEQ ID No. 158; SEQ ID No. 159;
SEQ ID No. 160; SEQ ID No. 161; SEQ ID No. 192;
SEQ ID No. 193; SEQ ID No. 196; SEQ ID No. 196;
SEQ ID No. 197; SEQ ID No. 198; SEQ ID No. 199;
SEQ ID No. 227; SEQ ID No. 229; SEQ ID No. 236;
SEQ ID No. 236; SEQ ID No. 239; SEQ ID No. 243;
SEQ ID No. 245; SEQ ID No. 264; SEQ ID No. 265;
SEQ ID No. 297; SEQ ID No. 331; SEQ ID No. 333;
SEQ ID No. 359; SEQ ID No. 360; SEQ ID No. 374;

SEQ ID No. 404; SEQ ID No. 405; SEQ ID No. 405;
SEQ ID No. 410; SEQ ID No. 415; SEQ ID No. 415;
SEQ ID No. 416; SEQ ID No. 417; SEQ ID No. 432;
SEQ ID No. 460; SEQ ID No. 461; SEQ ID No. 462;
SEQ ID No. 495; SEQ ID No. 513; SEQ ID No. 515;
SEQ ID No. 566; SEQ ID No. 566; SEQ ID No. 566;
SEQ ID No. 589; SEQ ID No. 613; SEQ ID No. 645;
SEQ ID No. 646; SEQ ID No. 647; SEQ ID No. 652;
SEQ ID No. 653; SEQ ID No. 654; SEQ ID No. 672;
SEQ ID No. 673; SEQ ID No. 674; SEQ ID No. 682;
SEQ ID No. 684; SEQ ID No. 692; SEQ ID No. 700;
SEQ ID No. 725; SEQ ID No. 801; SEQ ID No. 802;
SEQ ID No. 835; SEQ ID No. 836; SEQ ID No. 837;
SEQ ID No. 860; SEQ ID No. 861; SEQ ID No. 862;
SEQ ID No. 863; SEQ ID No. 869; SEQ ID No. 869;
SEQ ID No. 925; SEQ ID No. 964; SEQ ID No. 983 and one of their fragments.

36. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleotides, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 142; SEQ ID No. 142; SEQ ID No. 169;
SEQ ID No. 256; SEQ ID No. 268; SEQ ID No. 325;
SEQ ID No. 352; SEQ ID No. 366; SEQ ID No. 435;
SEQ ID No. 444; SEQ ID No. 528; SEQ ID No. 529;
SEQ ID No. 530; SEQ ID No. 548; SEQ ID No. 549;
SEQ ID No. 601; SEQ ID No. 602; SEQ ID No. 617;
SEQ ID No. 619; SEQ ID No. 644; SEQ ID No. 745;
SEQ ID No. 971; SEQ ID No. 972; SEQ ID No. 1023 and one of their representative fragments.

37. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of nucleic acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 5; SEQ ID No. 12; SEQ ID No. 82;

SEQ ID No. 96; SEQ ID No. 97; SEQ ID No. 98;
SEQ ID No. 99; SEQ ID No. 100; SEQ ID No. 105;
SEQ ID No. 118; SEQ ID No. 136; SEQ ID No. 137;
SEQ ID No. 163; SEQ ID No. 190; SEQ ID No. 204;
SEQ ID No. 259; SEQ ID No. 260; SEQ ID No. 262;
SEQ ID No. 290; SEQ ID No. 300; SEQ ID No. 301;
SEQ ID No. 302; SEQ ID No. 387; SEQ ID No. 427;
SEQ ID No. 434; SEQ ID No. 441; SEQ ID No. 444;
SEQ ID No. 471; SEQ ID No. 595; SEQ ID No. 596;
SEQ ID No. 597; SEQ ID No. 599; SEQ ID No. 600;
SEQ ID No. 605; SEQ ID No. 612; SEQ ID No. 624;
SEQ ID No. 625; SEQ ID No. 650; SEQ ID No. 657;
SEQ ID No. 658; SEQ ID No. 702; SEQ ID No. 703;
SEQ ID No. 704; SEQ ID No. 708; SEQ ID No. 719;
SEQ ID No. 766; SEQ ID No. 767; SEQ ID No. 775;
SEQ ID No. 779; SEQ ID No. 787; SEQ ID No. 788;
SEQ ID No. 794; SEQ ID No. 841; SEQ ID No. 842;
SEQ ID No. 883; SEQ ID No. 884; SEQ ID No. 907;
SEQ ID No. 918; SEQ ID No. 924; SEQ ID No. 928;
SEQ ID No. 929; SEQ ID No. 962; SEQ ID No. 962;
SEQ ID No. 963; SEQ ID No. 969; SEQ ID No. 970;
SEQ ID No. 975; SEQ ID No. 979; SEQ ID No. 995;
SEQ ID No. 1031; SEQ ID No. 1032 and one of their fragments.

38. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of amino acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 27; SEQ ID No. 41; SEQ ID No. 55;
SEQ ID No. 56; SEQ ID No. 57; SEQ ID No. 59;
SEQ ID No. 62; SEQ ID No. 63; SEQ ID No. 64;
SEQ ID No. 65; SEQ ID No. 119; SEQ ID No. 132;
SEQ ID No. 240; SEQ ID No. 241; SEQ ID No. 277;
SEQ ID No. 278; SEQ ID No. 279; SEQ ID No. 382;
SEQ ID No. 406; SEQ ID No. 428; SEQ ID No. 442;
SEQ ID No. 446; SEQ ID No. 447; SEQ ID No. 453;
SEQ ID No. 454; SEQ ID No. 541; SEQ ID No. 542;

SEQ ID No. 591; SEQ ID No. 608; SEQ ID No. 609;
SEQ ID No. 610; SEQ ID No. 618; SEQ ID No. 648;
SEQ ID No. 649; SEQ ID No. 660; SEQ ID No. 661;
SEQ ID No. 677; SEQ ID No. 717; SEQ ID No. 765;
SEQ ID No. 797; SEQ ID No. 871; SEQ ID No. 875;
SEQ ID No. 920; SEQ ID No. 922; SEQ ID No. 937;
SEQ ID No. 998; SEQ ID No. 1020; SEQ ID No. 1021;
SEQ ID No. 1034; SEQ ID No. 1044; SEQ ID No. 1046;
SEQ ID No. 1049 and one of their fragments.

39. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of polypeptides, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 21; SEQ ID No. 21; SEQ ID No. 22;
SEQ ID No. 23; SEQ ID No. 24; SEQ ID No. 25;
SEQ ID No. 26; SEQ ID No. 75; SEQ ID No. 84;
SEQ ID No. 84; SEQ ID No. 86; SEQ ID No. 92;
SEQ ID No. 133; SEQ ID No. 151; SEQ ID No. 152;
SEQ ID No. 157; SEQ ID No. 179; SEQ ID No. 209;
SEQ ID No. 307; SEQ ID No. 326; SEQ ID No. 343;
SEQ ID No. 344; SEQ ID No. 345; SEQ ID No. 371;
SEQ ID No. 429; SEQ ID No. 519; SEQ ID No. 557;
SEQ ID No. 586; SEQ ID No. 587; SEQ ID No. 630;
SEQ ID No. 656; SEQ ID No. 706; SEQ ID No. 707;
SEQ ID No. 730; SEQ ID No. 751; SEQ ID No. 752;
SEQ ID No. 786; SEQ ID No. 847; SEQ ID No. 885;
SEQ ID No. 923; SEQ ID No. 978; SEQ ID No. 1039;
SEQ ID No. 1048 and one of their fragments.

40. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the metabolism of fatty acids, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 4; SEQ ID No. 15; SEQ ID No. 16;
SEQ ID No. 141; SEQ ID No. 173; SEQ ID No. 205;
SEQ ID No. 205; SEQ ID No. 206; SEQ ID No. 207;

SEQ ID No. 208; SEQ ID No. 312; SEQ ID No. 355; SEQ ID No. 415; SEQ ID No. 550; SEQ ID No. 558; SEQ ID No. 560; SEQ ID No. 561; SEQ ID No. 574; SEQ ID No. 574; SEQ ID No. 577; SEQ ID No. 578; SEQ ID No. 590; SEQ ID No. 614; SEQ ID No. 772; SEQ ID No. 808; SEQ ID No. 809; SEQ ID No. 904; SEQ ID No. 905; SEQ ID No. 905; SEQ ID No. 933; SEQ ID No. 934; SEQ ID No. 934; SEQ ID No. 936 and one of their fragments.

41. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the synthesis of the wall, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 87; SEQ ID No. 196; SEQ ID No. 242; SEQ ID No. 269; SEQ ID No. 628; SEQ ID No. 629; SEQ ID No. 634; SEQ ID No. 635; SEQ ID No. 637; SEQ ID No. 638; SEQ ID No. 1019 and one of their fragments.

42. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the transcription, translation or maturation process, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 112; SEQ ID No. 113; SEQ ID No. 332; SEQ ID No. 212; SEQ ID No. 213; SEQ ID No. 350; SEQ ID No. 362; SEQ ID No. 363; SEQ ID No. 364; SEQ ID No. 407; SEQ ID No. 451; SEQ ID No. 546; SEQ ID No. 643; SEQ ID No. 744; SEQ ID No. 746; SEQ ID No. 833; SEQ ID No. 868; SEQ ID No. 981; SEQ ID No. 982; SEQ ID No. 1003; SEQ ID No. 1011; SEQ ID No. 1042 and one of their fragments.

43. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* ribosomal polypeptide or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 114; SEQ ID No. 115; SEQ ID No. 116;

SEQ ID No. 328; SEQ ID No. 361; SEQ ID No. 375; SEQ ID No. 445; SEQ ID No. 543; SEQ ID No. 584; SEQ ID No. 585; SEQ ID No. 743; SEQ ID No. 813; SEQ ID No. 941; SEQ ID No. 942; SEQ ID No. 944; SEQ ID No. 946; SEQ ID No. 947; SEQ ID No. 948; SEQ ID No. 950; SEQ ID No. 951; SEQ ID No. 952; SEQ ID No. 953; SEQ ID No. 954; SEQ ID No. 955; SEQ ID No. 955; SEQ ID No. 957; SEQ ID No. 958; SEQ ID No. 960; SEQ ID No. 961; SEQ ID No. 1040; SEQ ID No. 1041; SEQ ID No. 1043; SEQ ID No. 1063; SEQ ID No. 1064 and one of their fragments.

44. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* transport polypeptide or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:

SEQ ID No. 6; SEQ ID No. 50; SEQ ID No. 51; SEQ ID No. 80; SEQ ID No. 125; SEQ ID No. 126; SEQ ID No. 128; SEQ ID No. 129; SEQ ID No. 215; SEQ ID No. 246; SEQ ID No. 248; SEQ ID No. 249; SEQ ID No. 251; SEQ ID No. 252; SEQ ID No. 253; SEQ ID No. 255; SEQ ID No. 271; SEQ ID No. 275; SEQ ID No. 293; SEQ ID No. 309; SEQ ID No. 323; SEQ ID No. 324; SEQ ID No. 398; SEQ ID No. 401; SEQ ID No. 449; SEQ ID No. 511; SEQ ID No. 512; SEQ ID No. 564; SEQ ID No. 565; SEQ ID No. 667; SEQ ID No. 679; SEQ ID No. 680; SEQ ID No. 711; SEQ ID No. 712; SEQ ID No. 713; SEQ ID No. 714; SEQ ID No. 715; SEQ ID No. 730; SEQ ID No. 731; SEQ ID No. 736; SEQ ID No. 737; SEQ ID No. 738; SEQ ID No. 870; SEQ ID No. 908; SEQ ID No. 919; SEQ ID No. 977; SEQ ID No. 987; SEQ ID No. 988; SEQ ID No. 992; SEQ ID No. 993; SEQ ID No. 994; SEQ ID No. 1028; SEQ ID No. 1029 and one of their fragments.

45. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the virulence process, and in that it is chosen from

the polypeptides having the following sequences:
SEQ ID No. 20; SEQ ID No. 815; SEQ ID No. 816; SEQ ID No. 898; SEQ ID No. 1059; SEQ ID No. 1060 and one of their fragments.

46. Polypeptide according to one of Claims 5 to 8, characterized in that it is a *Chlamydia trachomatis* polypeptide or one of its fragments which is involved in the secretory system and/or which is secreted, and in that it is chosen from the polypeptides having the following sequences:
SEQ ID No. 758; SEQ ID No. 888; SEQ ID No. 889; SEQ ID No. 890; SEQ ID No. 891; SEQ ID No. 896; SEQ ID No. 897; SEQ ID No. 898 and one of their fragments.

47. Polypeptide according to one of Claims 5 to 8, characterized in that it is a polypeptide specific to *Chlamydiae* or one of its fragments, and in that it is chosen from the polypeptides having the following sequences:
SEQ ID No. 22; SEQ ID No. 29; SEQ ID No. 31;
SEQ ID No. 32; SEQ ID No. 34; SEQ ID No. 35;
SEQ ID No. 39; SEQ ID No. 40; SEQ ID No. 43;
SEQ ID No. 48; SEQ ID No. 49; SEQ ID No. 50;
SEQ ID No. 52; SEQ ID No. 53; SEQ ID No. 54;
SEQ ID No. 72; SEQ ID No. 77; SEQ ID No. 78;
SEQ ID No. 87; SEQ ID No. 90; SEQ ID No. 95;
SEQ ID No. 108; SEQ ID No. 110; SEQ ID No. 111;
SEQ ID No. 122; SEQ ID No. 123; SEQ ID No. 124;
SEQ ID No. 127; SEQ ID No. 138; SEQ ID No. 144;
SEQ ID No. 146; SEQ ID No. 153; SEQ ID No. 155;
SEQ ID No. 164; SEQ ID No. 166; SEQ ID No. 175;
SEQ ID No. 182; SEQ ID No. 184; SEQ ID No. 186;
SEQ ID No. 187; SEQ ID No. 188; SEQ ID No. 202;
SEQ ID No. 210; SEQ ID No. 247; SEQ ID No. 258;
SEQ ID No. 266; SEQ ID No. 267; SEQ ID No. 270;
SEQ ID No. 273; SEQ ID No. 274; SEQ ID No. 295;
SEQ ID No. 296; SEQ ID No. 305; SEQ ID No. 306;
SEQ ID No. 309; SEQ ID No. 318; SEQ ID No. 319;
SEQ ID No. 322; SEQ ID No. 326; SEQ ID No. 342;

SEQ ID No. 357; SEQ ID No. 376; SEQ ID No. 379;
SEQ ID No. 380; SEQ ID No. 388; SEQ ID No. 390;
SEQ ID No. 400; SEQ ID No. 431; SEQ ID No. 433;
SEQ ID No. 438; SEQ ID No. 443; SEQ ID No. 456;
SEQ ID No. 457; SEQ ID No. 458; SEQ ID No. 464;
SEQ ID No. 468; SEQ ID No. 470; SEQ ID No. 473;
SEQ ID No. 486; SEQ ID No. 489; SEQ ID No. 497;
SEQ ID No. 501; SEQ ID No. 503; SEQ ID No. 504;
SEQ ID No. 508; SEQ ID No. 512; SEQ ID No. 521;
SEQ ID No. 522; SEQ ID No. 523; SEQ ID No. 524;
SEQ ID No. 533; SEQ ID No. 535; SEQ ID No. 536;
SEQ ID No. 537; SEQ ID No. 538; SEQ ID No. 539;
SEQ ID No. 540; SEQ ID No. 554; SEQ ID No. 563;
SEQ ID No. 572; SEQ ID No. 579; SEQ ID No. 595;
SEQ ID No. 603; SEQ ID No. 604; SEQ ID No. 606;
SEQ ID No. 607; SEQ ID No. 615; SEQ ID No. 616;
SEQ ID No. 622; SEQ ID No. 641; SEQ ID No. 642;
SEQ ID No. 659; SEQ ID No. 668; SEQ ID No. 670;
SEQ ID No. 693; SEQ ID No. 695; SEQ ID No. 696;
SEQ ID No. 699; SEQ ID No. 703; SEQ ID No. 704;
SEQ ID No. 716; SEQ ID No. 726; SEQ ID No. 728;
SEQ ID No. 739; SEQ ID No. 742; SEQ ID No. 747;
SEQ ID No. 750; SEQ ID No. 751; SEQ ID No. 755;
SEQ ID No. 757; SEQ ID No. 759; SEQ ID No. 761;
SEQ ID No. 762; SEQ ID No. 763; SEQ ID No. 764;
SEQ ID No. 773; SEQ ID No. 780; SEQ ID No. 781;
SEQ ID No. 789; SEQ ID No. 800; SEQ ID No. 803;
SEQ ID No. 804; SEQ ID No. 818; SEQ ID No. 820;
SEQ ID No. 822; SEQ ID No. 823; SEQ ID No. 824;
SEQ ID No. 827; SEQ ID No. 828; SEQ ID No. 839;
SEQ ID No. 849; SEQ ID No. 850; SEQ ID No. 851;
SEQ ID No. 852; SEQ ID No. 855; SEQ ID No. 856;
SEQ ID No. 857; SEQ ID No. 858; SEQ ID No. 859;
SEQ ID No. 860; SEQ ID No. 861; SEQ ID No. 862;
SEQ ID No. 863; SEQ ID No. 865; SEQ ID No. 868;
SEQ ID No. 869; SEQ ID No. 870; SEQ ID No. 871;
SEQ ID No. 872; SEQ ID No. 873; SEQ ID No. 874;
SEQ ID No. 875; SEQ ID No. 877; SEQ ID No. 878;
SEQ ID No. 880; SEQ ID No. 882; SEQ ID No. 884;

SEQ ID No. 886; SEQ ID No. 893; SEQ ID No. 901; SEQ ID No. 906; SEQ ID No. 910; SEQ ID No. 912; SEQ ID No. 915; SEQ ID No. 916; SEQ ID No. 917; SEQ ID No. 926; SEQ ID No. 929; SEQ ID No. 933; SEQ ID No. 965; SEQ ID No. 967; SEQ ID No. 968; SEQ ID No. 984; SEQ ID No. 986; SEQ ID No. 989; SEQ ID No. 990; SEQ ID No. 996; SEQ ID No. 997; SEQ ID No. 1001; SEQ ID No. 1002; SEQ ID No. 1013; SEQ ID No. 1016; SEQ ID No. 1031; SEQ ID No. 1033; SEQ ID No. 1035; SEQ ID No. 1049; SEQ ID No. 1051; SEQ ID No. 1052; SEQ ID No. 1054; SEQ ID No. 1056; SEQ ID No. 1057; SEQ ID No. 1058; SEQ ID No. 1062; SEQ ID No. 1070; SEQ ID No. 1071; SEQ ID No. 1073 and one of their fragments.

48. Nucleotide sequence according to one of Claims 1 to 4, and 9 to 28, and/or a polypeptide sequence according to one of Claims 5 to 8, and 29 to 47, characterized in that the said sequence(s) are recorded on a recording medium whose type and nature facilitate the reading, the analysis and/or the exploitation of the said sequence(s).

49. Nucleotide sequence which can be used as a primer, characterized in that the said sequence is chosen from the nucleotide sequences according to one of Claims 2 to 4, and 9 to 28.

50. Nucleotide sequence which can be used as a probe, characterized in that the said sequence is chosen from the nucleotide sequences according to one of Claims 2 to 4, and 9 to 28.

51. Nucleotide sequence according to either of Claims 49 and 50, characterized in that it is labelled with a radioactive compound or with a nonradioactive compound.

52. Nucleotide sequence according to one of Claims 49 to 51, characterized in that it is covalently or noncovalently immobilized on a support.

53. Nucleotide sequence according to one of Claims 49 to 52, characterized in that it is immobilized on a support of a DNA chip.

54. Nucleotide sequence according to one of Claims 49 to 53, for the detection and/or the amplification of nucleic sequences.

55. DNA chip, characterized in that it contains at least one nucleotide sequence according to Claim 53.

56. DNA chip according to Claim 55, characterized in that it contains, in addition, at least one nucleotide sequence of a microorganism different from *Chlamydia trachomatis,* immobilized on the support of the said chip.

57. DNA chip according to Claim 56, characterized in that the different microorganism is chosen from a microorganism associated with *Chlamydia trachomatis,* a bacterium of the *Chlamydia* family, and a variant of *Chlamydia trachomatis.*

58. Cloning and/or expression vector, characterized in that it contains a nucleotide sequence according to one of Claims 2 to 5, and 9 to 28.

59. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to one of Claims 10 to 12.

60. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to one of Claims 13 to 15.

61. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to one of Claims 16 to 21.

62. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to one of Claims 22 to 25.

63. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to Claim 26.

64. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to Claim 27.

65. Vector according to Claim 58, characterized in that it contains a nucleotide sequence according to Claim 28.

66. Host cell, characterized in that it is transformed with a vector according to one of Claims 58 to 65.

67. Host cell according to Claim 66, characterized in that it is a bacterium belonging to the *Chlamydia* family.

68. Host cell according to Claim 67, characterized in that it is a bacterium belonging to the species *Chlamydia trachomatis*.

69. Host cell according to Claim 66, characterized in that it is a microorganism associated with the species *Chlamydia trachomatis*.

70. Animal, except humans, comprising a transformed cell according to one of Claims 66 to 69.

71. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 58, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

72. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 59, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

73. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 60, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

74. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 61, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

75. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 62, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

76. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 63, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

77. Method of preparing a polypeptide, charac-

terized in that it uses a vector according to Claim 64, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

78. Method of preparing a polypeptide, characterized in that it uses a vector according to Claim 65, a cell transformed with the said vector and/or an animal comprising the said transformed cell.

79. Recombinant polypeptide capable of being obtained by a method according to Claim 71.

80. Recombinant polypeptide capable of being obtained by a method according to Claim 72.

81. Recombinant polypeptide capable of being obtained by a method according to Claim 73.

82. Recombinant polypeptide capable of being obtained by a method according to Claim 74.

83. Recombinant polypeptide capable of being obtained by a method according to Claim 75.

84. Recombinant polypeptide capable of being obtained by a method according to Claim 76.

85. Recombinant polypeptide capable of being obtained by a method according to Claim 77.

86. Recombinant polypeptide capable of being obtained by a method according to Claim 78.

87. Method of preparing a synthetic polypeptide, characterized in that it uses an amino acid sequence of a polypeptide according to one of Claims 5 to 8, 29 to 47, and 79 to 86.

88. Synthetic polypeptide obtained by a method according to Claim 87.

89. Hybrid polypeptide, characterized in that it comprises at least the sequence of a polypeptide according to one of Claims 5 to 8, 29 to 47, and 79 to 86, and 88, and a sequence of a polypeptide capable of eliciting an immune response in humans or animals.

90. Hybrid polypeptide according to Claim 89, characterized in that it comprises at least the sequence of a polypeptide according to one of Claims 29 to 31, and 80, and a sequence of a polypeptide capable of eliciting an immune response in humans or animals.

91. Hybrid polypeptide according to either of Claims 89 and 90, characterized in that the polypeptide capable of eliciting an immune response contains at least one antigenic determinant capable of eliciting a humoral and/or cellular response.

92. Nucleotide sequence encoding a hybrid polypeptide according to one of Claims 89 to 91.

93. Vector, characterized in that it contains a nucleotide sequence according to Claim 92.

94. Hybrid polypeptide according to one of Claims 89 to 91, characterized in that it is a recombinant polypeptide obtained using a vector according to Claim 93.

95. Method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it comprises the following steps:

a) bringing the biological sample into contact with a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88;

b) detecting the antigen-antibody complex which may be formed.

96. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88;

b) where appropriate, the reagents for constituting the medium appropriate for the immunological reaction;

c) the reagents allowing the detection of the antigen-antibody complexes which may be formed between the polypeptide(s) of the invention and the antibodies;

d) where appropriate, a reference biological sample (negative control) free of antibodies recognized by the said polypeptide;

e) where appropriate, a reference biological sample

(positive control) containing a predetermined quantity of antibodies recognized by the said polypeptide.

97. Mono- or polyclonal antibodies, their fragments, or chimeric antibodies, characterized in that they are capable of specifically recognizing a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88.

98. Antibody according to Claim 97, characterized in that it is a labelled antibody.

99. Method for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it comprises the following steps:

a) bringing the biological sample into contact with an antibody according to either of Claims 97 and 98;

b) detecting the antigen-antibody complex formed.

100. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a polyclonal or monoclonal antibody according to either of Claims 97 and 98;

b) where appropriate, the reagents for constituting the medium appropriate for the immunological reaction;

c) the reagents allowing the detection of the antigen-antibody complexes produced by the immunological reaction.

101. Polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88, or antibodies according to either of Claims 97 and 98, characterized in that it is immobilized on a support, in particular of a protein chip.

102. Protein chip, characterized in that it contains at least one polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88, or at least one antibody according to either of Claims 97 and 98, immobilized on the support of the said chip.

103. Protein chip according to Claim 102, characterized in that it contains, in addition, at least one polypeptide of a microorganism different from *Chlamydia trachomatis* or at least one antibody directed against a compound of a microorganism different from *Chlamydia trachomatis*, immobilized on the support of the said chip.

104. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises a protein chip according to either of Claims 102 and 103.

105. Kit or set for the detection and/or the identification of a microorganism, characterized in that it comprises a protein chip according to Claim 103.

106. Method of detection and/or of identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism in a biological sample, characterized in that it uses a nucleotide sequence according to one of Claims 1 to 4, 9 to 28, 49 to 54.

107. Method according to Claim 106, characterized in that it comprises the following steps:

a) where appropriate, isolation of the DNA from the biological sample to be analysed, or optionally production of a cDNA from the RNA in the biological sample;

b) specific amplification of the DNA of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism with the aid of at least one primer according to one of Claims 49 to 54;

c) detection of the amplification products.

108. Method according to Claim 106, characterized in that it comprises the following steps:

a) bringing a nucleotide probe according to one of Claims 50 to 54 into contact with a biological sample, the DNA contained in the biological sample having, where appropriate, been previously made accessible to

hybridization, under conditions allowing the hybridization of the probe to the DNA of a bacterium belonging to the species *Chlamydia trachomatis* or to an associated microorganism;

b) detecting the hybrid which may be formed between the nucleotide probe and the DNA in the biological sample.

109. Method according to Claim 106, characterized in that it comprises the following steps:

a) bringing a nucleotide probe immobilized on a support according to Claim 52 into contact with a biological sample, the DNA in the sample having, where appropriate, been previously made accessible to hybridization, under conditions allowing the hybridization of the probe to the DNA of a bacterium belonging to the species *Chlamydia trachomatis* or to an associated microorganism;

b) bringing the hybrid formed between the nucleotide probe immobilized on a support and the DNA contained in the biological sample, where appropriate after removal of the DNA in the biological sample which has not hybridized with the probe, into contact with a labelled nucleotide probe according to Claim 51;

c) detecting the new hybrid formed in step b).

110. Method according to Claim 109, characterized in that, prior to step a), the DNA in the biological sample, or the cDNA which may have been obtained by reverse transcription of the RNA in the sample, is amplified with the aid of at least one primer according to one of Claims 49 to 54.

111. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a nucleotide probe according to one of Claims 50 to 54;

b) where appropriate, the reagents necessary for carrying out a hybridization reaction;

c) where appropriate, at least one primer according

to one of Claims 49 to 54 as well as the reagents necessary for a DNA amplification reaction.

112. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) a nucleotide probe, called capture probe, according to Claim 52;

b) an oligonucleotide probe, called detection probe, according to Claim 51;

c) where appropriate, at least one primer according to one of Claims 49 to 54 as well as the reagents necessary for a DNA amplification reaction.

113. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises the following components:

a) at least one primer according to one of Claims 49 to 54;

b) where appropriate, the reagents necessary for carrying out a DNA amplification reaction;

c) where appropriate, a component which makes it possible to check the sequence of the amplified fragment, more particularly an oligonucleotide probe according to one of Claims 50 to 54.

114. Kit or set for the detection and/or the identification of bacteria belonging to the species *Chlamydia trachomatis* or to an associated microorganism, characterized in that it comprises a DNA chip according to Claim 55.

115. Kit or set for the detection and/or the identification of a microorganism, characterized in that it comprises a DNA chip according to either of Claims 56 and 57.

116. Method or kit or set according to one of Claims 95, 96, 99, 100 and 104 to 115, for the detection and/or the identification of bacteria

belonging to the species *Chlamydia trachomatis*, characterized in that the said primer and/or the said probe are chosen from the nucleotide sequences according to one of Claims 2 to 4, 9 to 28, and 49 to 54 specific to the species *Chlamydia trachomatis*, in that the said polypeptides are chosen from the polypeptides according to one of Claims 5 to 8, 29 to 47, 79 to 86, 88 and 101 specific to the species *Chlamydia trachomatis* and in that the said antibodies are chosen from the antibodies according to one of Claims 97 to 98 directed against the polypeptides chosen from the polypeptides according to one of Claims 5 to 8, 29 to 47, 79 to 86, 88 to 101 specific to the species *Chlamydia trachomatis*.

117. Method or kit or set according to Claim 116, characterized in that the said primer and/or the said probe are chosen from the nucleotide sequences according to Claim 28, in that the said polypeptides are chosen from the polypeptides according to either of Claims 47 and 86, and in that the said antibodies are chosen from the antibodies according to either of Claims 97 and 98 directed against the polypeptides according to either of Claims 47 and 86.

118. Method or kit or set according to one of Claims 95, 96, 99, 100 and 104 to 117, for the diagnosis of predispositions to genital diseases, which are induced or worsened by a *Chlamydia trachomatis* infection.

119. Method or kit or set according to one of Claims 95, 96, 99, 100 and 104 to 117, for the diagnosis of predispositions to, or of conditions caused by, eye diseases induced or worsened by a *Chlamydia trachomatis* infection.

120. Method or kit or set according to one of Claims 95, 96, 99, 100 and 104 to 117, for the diagnosis of predispositions to, or of conditions caused by, systemic diseases, especially of the lymphatic system, which are induced or worsened by the *Chlamydia trachomatis* infection.

121. Use of a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86 and 88, a transformed cell according to one of Claims 66 to 69 and/or an animal according to Claim 70, for the biosynthesis or the biodegradation of a compound of interest.

122. Method of biosynthesis or of biodegradation of a compound of interest, characaterized in that it uses a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88, a transformed cell according to one of Claims 66 to 69 and/or an animal according to Claim 70.

123. Use of a nucleotide sequence according to one of Claims 2 to 4, and 9 to 28, of a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88, of an antibody according to either of Claims 97 and 98, of a cell according to one of Claims 66 to 69, and/or of a transformed animal according to Claim 70, for the selection of an organic or inorganic compound capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or of modifying the cellular replication of eukaryotic or prokaryotic cells or capable of inducing, inhibiting or worsening in an animal or human organism the pathologies linked to an infection by *Chlamydia trachomatis* or by an associated microorganism.

124. Method of selecting a compound capable of binding to a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, and 88, capable of binding to a nucleotide sequence according to one of Claims 2 to 4, and 9 to 28, or capable of recognizing an antibody according to either of Claims 97 and 98, and/or capable of modulating, regulating, inducing or inhibiting the expression of genes, and/or of modifying the cellular replication of eukaryotic or prokaryotic cells, or capable of inducing, inhibiting or worsening, in an animal or human organism, the pathologies linked to an infection by *Chlamydia trachomatis*, characterized in that it comprises the following steps:

a) bringing the said compound into contact with the

said polypeptide, the said nucleotide sequence, with a transformed cell according to one of Claims 66 to 69 and/or administering the said compound to a transformed animal according to Claim 70;

b) determining the capacity of the said compound to bind with the said polypeptide or the said nucleotide sequence, or to modulate, regulate, induce or inhibit the expression of genes, or to modulate growth or cellular replication, or to induce, inhibit or worsen in the said animal or human organism, the pathologies linked to an infection by *Chlamydia trachomatis* or by an associated microorganism.

125. Compound capable of being selected by a method according to Claim 124.

126. Pharmaceutical composition comprising a compound chosen from the following compounds:

a) a nucleotide sequence according to one of Claims 2 to 4, 9 to 28;

b) a polypeptide according to one of Claims 5 to 8, 29 to 47, 79 to 86, 88 to 91, and 94;

c) a vector according to one of Claims 58 to 65, and 93;

d) an antibody according to Claim 97; and

e) a compound according to Claim 125.

127. Composition according to Claim 126, optionally in combination with a pharmaceutically acceptable vehicle.

128. Pharmaceutical composition according to either of Claims 126 and 127 for the prevention or the treatment of an infection by a bacterium belonging to the species *Chlamydia trachomatis* or by an associated microorganism.

129. Vaccine composition, characterized in that it comprises one or more polypeptides according to one of Claims 5 to 8, 29 to 47, 79 to 86, 88, and/or one or more hybrid polypeptides according to one of Claims 89 to 91 and 94.

130. Use of a cell according to one of Claims 66 to 69, for the preparation of a vaccine composition.

131. Vaccine composition, characterized in that it contains a vector according to one of Claims 58 to 65, and 93, and/or a cell according to one of Claims 66 to 69.

132. Vaccine composition according to either of Claims 129 and 131, for the prevention or the treatment of an infection by a bacterium belonging to the species *Chlamydia trachomatis* or by an associated microorganism.

133. Vaccine composition according to one of Claims 129, 131 to 132, in combination with a pharmaceutically acceptable vehicle and, where appropriate, one or more appropriate immunity adjuvants.

134. Use of a composition according to one of Claims 126 to 129, and 131 to 133, for the preparation of a medicament intended for the treatment and/or the prevention of genital diseases which are induced or worsened by *Chlamydia trachomatis*.

135. Use of a composition according to one of Claims 126 to 129, and 131 to 133, for the preparation of a medicament intended for the treatment and/or the prevention of eye diseases which are induced or worsened by *Chlamydia trachomatis*.

136. Use of a composition according to one of Claims 126 to 129, and 131 to 133, for the preparation of a medicament intended for the treatment and/or the prevention of systemic diseases which are induced or worsened by the presence of *Chlamydia trachomatis*.

137. Use according to Claim 136, characterized in that the systemic disease affects the lymphatic system.

PATENT OF INVENTION

=-=-=-=-=-=-=-=-=-=-=

TITLE

CHLAMYDIA TRACHOMATIS GENOMIC SEQUENCE AND POLYPEPTIDES, FRAGMENTS THEREOF AND USES THEREOF, IN PARTICULAR FOR THE DIAGNOSIS, PREVENTION AND TREATMENT OF INFECTION

DESCRIPTIVE ABSTRACT

The subject of the invention is the genomic sequence and the nucleotide sequences encoding polypeptides of *Chlamydia trachomatis*, such as cellular envelope polypeptides, which are secreted or specific, or which are involved in metabolism, in the replication process or in virulence, as well as vectors including the said sequences and cells or animals transformed with these vectors. The invention also relates to methods of detecting these nucleic acids or polypeptides and kits for diagnosing *Chlamydia trachomatis* infection. The invention also relates to a method of selecting compounds capable of modulating bacterial infection and a method for the biosynthesis or biodegradation of molecules of interest using the said nucleotide sequences or the said polypeptides. The invention finally comprises, pharmaceutical, in particular vaccine, compositions for the prevention and/or the treatment of bacterial, in particular *Chlamydia trachomatis*, infections.